

UNDERSTANDING THE TRAINING AND TRANSFER EFFECTS IN N-BACK TRAINING

A Dissertation
Presented to
The Academic Faculty

by

Shivangi Jain

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
SCHOOL OF PSYCHOLOGY

Georgia Institute of Technology
AUGUST 2018

COPYRIGHT © 2018 BY SHIVANGI JAIN

UNDERSTANDING THE TRAINING AND TRANSFER EFFECTS IN N-BACK TRAINING

Approved by:

Dr. Paul Verhaeghen, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Julia Karbach
Faculty of Psychology
Universität Koblenz-Landau

Dr. Christopher Hertzog
School of Psychology
Georgia Institute of Technology

Dr. Eric Schumacher
School of Psychology
Georgia Institute of Technology

Dr. Susanne Jaeggi
School of Education
University of California-Irvine

Date Approved: May 8, 2018

To Papa

ACKNOWLEDGEMENTS

I would like to formally acknowledge and thank all the wonderful people who have been with me and have supported me through everything, including my time spent in graduate school. A sincere thank you to my parents- Poonam Jain and Vinod Kumar Jain, without whose unconditional support and trust, I would have never been able to travel half way across the world and pursue a career in psychological research. Their faith in me helped me grow into an independent and strong willed person, for which I'm ever grateful. A special thanks to my sisters- Urvashi Jain and Meenakshi Jain for picking up the slack and filling in for me while I was busy with graduate school and cheering me on, which made this journey a lot easier.

A heartfelt thanks to my advisor extraordinaire Paul Verhaeghen, who patiently guided me and went above and beyond to help me in any way he could. I owe all my academic accomplishments to him. His witty, unassuming and compassionate nature made my graduate school experience as less intense as it could be. A special thank you to Shelley Aikman and Bodhi Aikman Verhaeghen for their warmth and support, as well as to Paul and Shelley's fur kids- Dogen, Enzo and Zoe, who provided much needed pet therapy in times of stress.

An earnest thank you to my husband Arnab Roy, for providing unconditional love and support, for patiently dealing with my fickle mood, for showing understanding in my times of stress, for listening attentively to all my rants, and for completing all my unreasonable requests- be it to make coffee at 3am or to sleep on the couch while I was

writing my dissertation. Without him by my side, graduate school would have been more challenging.

A special shout out to my amazing lab mates and dear friends- Didem Pehlivanoglu and John Price. They were always there to brain storm ideas, give honest feedbacks and provide emotional support. Extra special thanks to Didem for bringing me goodies while I was writing my dissertation and to Divya Raghavan for being an understanding and supportive roommate. I would also like to thank Jay Prakash Patel for taking out time from his busy schedule and helping me with the coding of the experiments.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
LIST OF TABLES	viii
LIST OF FIGURES	ix
SUMMARY	xii
CHAPTER 1. Introduction	1
1.1 Relation between Working Memory and Fluid Intelligence	2
1.1.1 Executive Processes Linking Working Memory and Fluid Intelligence	4
1.2 Process-Based Training	7
1.2.1 Process-Based Working Memory Training in Young Adults	9
1.3 Meta-Analyses on Working Memory Training	12
1.4 Understanding the Underlying Mechanism in Working Memory Training:	
Present study	13
1.4.1 N-Back Task Decomposition	14
1.4.2 Study Details	15
CHAPTER 2. Method	19
2.1 Participants	19
2.2 Training Task	19
2.3 Pre-Post Assessments	21
2.3.1 Modified N-Back Task	21
2.3.2 Updating	21
2.3.3 Task Switching	22
2.3.4 Inhibition	23
2.3.5 Focus Switching	24
2.3.6 Fluid Intelligence	26
2.3.7 Theories of Cognitive Abilities (TOCA; Dweck, 1999)	26
CHAPTER 3. Results	28
3.1 Training Tasks	28
3.2 Pre and Post Training Assessment of the N-Back Component Processes	30
3.2.1 Trained Tasks	30
3.3 Pre and Post Training Transfer Tasks Assessment Error! Bookmark not defined.	
3.3.1 Fluid Intelligence	48
3.3.2 Updating	49
3.3.3 Task Switching	50
3.3.4 Inhibition	52
3.3.5 Focus Switching	53
3.4 Relation Between Change in N-Back Components and Transfer Measures	54
3.5 Strategy Questionnaire	55

CHAPTER 4. Discussion	67
4.1 Learning Curves for the Trained Tasks	67
4.2 Performance on the Trained Tasks	69
4.3 Effect of Training on the N-Back Components	72
4.4 Transfer Effects of the N-Back Training	74
4.5 N-Back Components and Transfer Measures	76
4.6 Relation Between Strategies and Performance	77
4.7 Limitations and Future Direction	78
4.8 Summary	80
REFERENCES	81

LIST OF TABLES

Table 1	- Last N level reached after each training session for Forward-no-updating (FWNU), Random-no-updating (RNU) and Forward-updating (FWU) conditions in the experimental group (exp) and the active control group (active).	28
Table 2	- Fit parameters after fitting negative exponential function ($y = a - b \cdot \exp(-c \cdot x)$) to the training data of Forward-no-updating (FWNU), Random-no-updating (RNU) and Forward-updating (FWU) conditions in experimental (exp) and active control groups (active)	29
Table 3	- Correlation between the different transfer measures at pre training.	46
Table 4	- Factor loadings based on principal component analysis with oblimin rotation for the transfer tasks at pre training.	46
Table 5	- Means and standard deviations for the different transfer measures at pre and post training for the experimental and the active control group.	47
Table 6	- Correlation between change in the N-Back components, TOCA and change in the transfer measures.	55
Table 7	- Correlation between the different strategies used during training sessions 9/10 and the average level of N reached during the same sessions.	65

LIST OF FIGURES

Figure 1	- Example 3-Back tasks: FWNU -- forward access without updating; FWUP -- forward access with updating; RNNU -- random access without updating (Price et al., 2014)	17
Figure 2	- Negative exponential function fitted to the training data of Forward-no-updating (FWNU), Random-no-updating (RNU) and Forward-updating (FWU) conditions in the experimental (exp) and active control groups (active). The circles represent the actual N-Back level reached after each training session. Solid lines represent the fitted curve.	30
Figure 3	- Figure shows the RT data for the forward-no-updating task at pre and post training assessment for the active control (active) and the experimental (exp) group. There was a significant main effect of time and N-Back level. Error bars denote the standard error of the mean.	32
Figure 4	- Figure shows the RT data for the non-switch trials of the random-no-updating task at pre and post training assessment for the active control (active) and the experimental (exp) group. Main effect of time and N-Back level was significant. In addition, interaction between time and N-Back level was also significant. Error bars denote the standard error of the mean.	34
Figure 5	- Figure shows the RT data for the switch trials of the random-no-updating task at pre and post training assessment for the active control (active) and the experimental (exp) group. Significant effect of time, N-Back level and interaction between time and N-Back level was present. Error bars denote the standard error of the mean.	36
Figure 6	- Figure shows the RT data for the forward-updating task at pre and post training assessment for the active control (active) and the experimental (exp) group. There was a significant main effect of time, N-Back level and group, as well as significant interaction between time and group, and between time and N-Back level. Error bars denote the standard error of the mean.	38
Figure 7	- Figure shows the focus switching direct retrieval component at pre and post training assessment for the active control (active) and the experimental (exp) group. There were no significant main effects and interactions present. Error bars denote the	40

standard error of the mean.

Figure 8	- Figure shows the focus switching requiring search component at pre and post training assessment for the active control (active) and the experimental (exp) group. It is the difference RT between switch and non-switch trials calculated from the random-no-updating task. Significant main effect of N-Back level was present for both the groups. Error bars denote the standard error of the mean.	41
Figure 9	- Figure shows the updating component at pre and post training assessment for the active control (active) and the experimental (exp) group. It is the difference RT between forward-updating and the forward-no-updating tasks. There was a significant main effect of time, group and N-Back level. In addition time x group and time x N-Back level interactions were also significant. Error bars denote the standard error of the mean.	43
Figure 10	- Figure shows the percent accuracy score on the fluid intelligence measures before (pre) and after (post) training for the active control (active) and the experimental group (exp). Panel A- BOMAT; Panel B- RAPM. Significant effect of group was present for both the tasks. Error bars denote the standard error of the mean.	49
Figure 11	- Figure shows the percent accuracy score on the updating measures before (pre) and after (post) training for the active control (active) and the experimental group (exp). Panel A- Keep track; Panel B- Letter memory. Significant effect of time and time by group interaction was present for the letter memory task, whereas, only the main effect of time was significant for the keep track task. Error bars denote the standard error of the mean.	50
Figure 12	- Figure shows the switching cost on the task switching measures before (pre) and after (post) training for the active control (active) and the experimental group (exp). Panel A- Plus minus; Panel B- Number letter. For the plus minus task, no main effects or interaction effects were significant. For the number letter task, main effects of time and group were present. Error bars denote the standard error of the mean.	51
Figure 13	- Figure shows the before (pre) and after (post) training scores for the active control (active) and the experimental group on the measures of inhibition. Panel A shows the difference RT between incongruent and neutral trials in the Stroop task; Panel B shows the percent accuracy for the antisaccade task; Panel C shows the percent accuracy for the stop signal task. Significant	53

effect of time was present for the Stroop and antisaccade tasks, no main effects and interactions were significant for the stop signal task. Error bars denote the standard error of the mean.

- | | | |
|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 14 | <p>- Figure shows the switching cost on the focus switching measures before (pre) and after (post) training for the active control (active) and the experimental group (exp). Panel A- Garavan; Panel B- Spatial focus switching. Main effect of time was significant for the Garavan task, no significant effects were present for the spatial focus switching task. Error bars denote the standard error of the mean.</p> | 54 |
| Figure 15 | <p>- The figure shows the use of nine different strategies reported by the experimental group (exp) and the active control (active) group to perform the forward-no-updating (FWNU), random-no-updating (RNU) and the forward-updating (FWU) tasks across the ten training sessions. S1- Sessions 1/3, S2- Sessions 5/7, S3- Sessions 9/10; Panel A- chunking, Panel B- memorizing, Panel C- visualizing, Panel D- pattern recognition, Panel E- episodic association, Panel F- recollection, Panel G- speed, Panel H- anchoring, Panel I- rhythm identification. Error bars denote the standard error of the mean.</p> | 63 |

SUMMARY

Current research looking at the effect of working memory training on constructs such as fluid intelligence has generated mixed findings. Some researchers have found that training participants on working memory tasks leads to an improvement on fluid intelligence scores, others have failed to find this effect. To reconcile these different findings, there is a need to understand the underlying mechanism of the transfer effect. In this study, a modified N-Back task was decomposed into its component processes, namely updating, focus switching one-step retrieval, and focus switching requiring search; the effect of training on each of the components was examined. Since updating has been found to be associated with both working memory (e.g., Miyake et al., 2000) and fluid intelligence (e.g., Friedman et al., 2006), the study specifically looked at the role of the updating component in eliciting transfer to other cognitive control processes (task switching and inhibition) as well as measures of fluid intelligence. The study employed two groups of participants—experimental and active control, which were trained for 10 hours over a period of two weeks and assessed on the transfer measures before and after training. The experimental group was trained on all the components of the modified N-Back task, whereas the active control group was trained on all components, except updating. If updating were the crucial link between training and transfer effects, the two groups should have shown differential effects on the transfer measures. However, this hypothesis was not supported. Training in the updating aspect of the N-Back task did not generalize to other cognitive control processes implicated in working memory nor did it lead to transfer to measures of fluid intelligence. However,

we did find that training effects are different on the different working memory components. The updating component is more malleable than the focus switching requiring search and focus switching direct retrieval. Thus working memory training protocols targeting the updating component might be more effective than the ones which don't include it.

CHAPTER 1. INTRODUCTION

Fluid intelligence is the ability to reason and solve problems in novel contexts. It is involved in many important cognitive abilities such as academic achievement, learning, problem solving, reading comprehension and reasoning. The idea that we can increase intelligence by training is enticing and has been researched for a long time. Recent focus in this area has been to see if cognitive tasks that are related to intelligence can boost scores on the construct by simply practicing those tasks over and over again. This type of training is termed process-based training and has shown improvement on the trained task and in some cases transfer to untrained tasks, which might share some of the same underlying basic cognitive processes.

Working memory is one of the constructs that have been trained to see the transfer effects to other cognitive abilities (e.g. Chein & Morrison, 2010; Jaeggi, Buschkuhl, Jonides, & Perrig, 2008), including fluid intelligence. The research so far has generated mixed results (e.g. Jaeggi et al., 2008; Morrison & Chein, 2011; Redick et al., 2013). The findings show that the performance on the trained task improves with training, but only a few studies have found transfer to untrained tasks measuring the trained construct (near-transfer; e.g. Dahlin, Neely, Larsson, Bäckman, & Nyberg, 2008; Morrison & Chein, 2011) or an untrained construct (far-transfer; e.g. Chein & Morrison, 2010; Jaeggi et al., 2008). Others have failed to find such transfer effects (e.g. Chooi & Thompson, 2012; Harrison et al., 2013; Redick et al., 2013). Most of these studies differ in their training protocol and transfer measures but even studies that have used similar training protocols have found varied transfer effects.

One way to understand these conflicting findings is to investigate the underlying mechanism that links practice effects in working memory with potential transfer effects on other cognitive tasks, including fluid intelligence. A common assumption in these training studies is that training participants on one task of working memory trains the whole construct, and that this change in working memory efficacy is the locus of any transfer effects. However, that may not be the case. Working memory tasks consist of many constituent processes, and it is possible that training affects each component process differently, and that changes in some of these components are more likely to lead to transfer effects. Thus, the inconsistent findings in literature may be an artifact of using different kinds of working memory training tasks, which differ in their constituent processes.

1.1 Relation between Working Memory and Fluid Intelligence

Working memory is the ability to temporarily store and manipulate information for successfully carrying out the task at hand. Working memory training is one of the preferred training regimens to examine transfer to measures of fluid intelligence in current research. Even though working memory and fluid intelligence are two distinct psychological constructs, research has shown that working memory supports cognitive functions such as logical reasoning and problem solving and is also strongly related to measures of fluid intelligence (Ackerman, Beier, & Boyle, 2005; Conway, Cowan, Bunting, Theriault, & Minkoff, 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Friedman, Miyake, Corley, Young, DeFries, & Hewitt, 2006; Kane, Hambrick, & Conway, 2005). Importantly, working memory measures that only tap storage (i.e., short-term memory, STM, tasks) are not related to fluid intelligence, as found in a seminal

individual-differences study by Engle et al. (1999).

Conway et al. (2002) extended the work done by Engle et al. (1999) by examining if the relation between working memory capacity; as measured by tasks involving both storage and processing, such as operation span and reading span) and fluid intelligence is mediated by processing speed as argued by some researchers (Fry & Hale, 1996; Kail & Salthouse, 1994; Salthouse, 1996). According to this argument, the faster the rate of processing, the greater the amount of information that can be processed in one unit of time. Thus, an individual's higher working memory capacity may be causally linked to higher processing speed (Kail & Salthouse, 1994; Salthouse, 1996). Conway et al. found that the SEM model with the best fit related working memory capacity directly to fluid intelligence, and that neither STM capacity or processing speed were a good predictor of fluid intelligence. Given that only working memory capacity tasks necessitate attentional control, the researchers suggested that controlled attention and/or strategic processes underlie performance on working memory span tasks as well as tests of fluid intelligence.

Kane et al. (2004) addressed two further questions: First, whether working memory capacity is a general or domain-specific construct and, second, how strongly it relates to both domain-general and domain-specific aspects of fluid reasoning. Confirmatory factor analyses indicated that working memory tasks largely reflect a domain-general factor: Verbal and spatial working memory capacity factors shared 70%–85% of their variance. The results also showed that working memory capacity is a strong predictor of fluid intelligence ($r = .60$) and a weak predictor of domain-specific reasoning.

Numerous studies have consistently reported substantial positive correlations between measures of fluid intelligence and working memory but the correlation value ($r = .48$ to $.85$) is not always agreed upon. Using latent variable analysis technique, these relationships sometimes even approach a perfect correlation suggesting that fluid intelligence and working memory might be a single construct (e.g. Buehner, Krumm & Pick, 2005; Colom, Rebollo, Palacios, Juan-Espinoso & Kyllonen 2004). To shed more light on this issue, Ackerman et al. (2005) conducted a meta-analysis that revealed an average correlation of $.48$ between fluid intelligence and working memory, indicating that intelligence and working memory are fairly strongly correlated but are still distinguishable at the manifest level. The correlation is higher at the latent level: Two meta-analyses that used a latent variable approach estimate the correlation at $.72$ (Kane et al., 2005), or $.85$ (Oberauer, Schulze, Wilhelm, & Süß, 2005). Working memory and intelligence thus can be seen as highly correlated constructs that, however, are not identical. Oberauer et al. state that working memory capacity should be regarded as an explanatory construct for intellectual abilities. Working memory capacity is a very strong predictor of reasoning ability and fluid intelligence. This strong relation has been the logic behind training protocols that have used working memory tasks in the hope that its effects would transfer to fluid intelligence measures.

1.1.1 Executive Processes Linking Working Memory and Fluid Intelligence

Working memory is not a unitary process but is determined by the interaction between several process components. In working memory information is actively maintained and manipulated in order to successfully carry out the task at hand. Ericksson, Vogel, Lansner, Bergstrom and Nyberg (2015) state some of the component processes

involved in working memory. For information to be properly encoded in working memory, selective attention acts on the perceptual information present or long-term stored representations; additionally to successfully maintain information, inhibition acts on the task-irrelevant stimuli. To prevent information decay, the encoded information is rehearsed and actively maintained by sustained attention. If the information to be maintained does not fit in the focus of attention (the capacity limited store of working memory), then an added rehearsal process in addition to active maintenance is required to prevent the information from decaying. At the retrieval phase selective attention and pattern completion processes are involved. The information maintained in working memory is matched to the perceptual input available to ensure successful task-relevant retrieval. If information held in working memory needs to be manipulated according to the task requirement, it is either done by performing manipulation operations (e.g., arithmetic operations in a computation-span working memory task) or by updating the current content (e.g., in an N-Back task or running span task). Task set, prospective planning, and other cognitive control operations are also involved in working memory.

Broadly speaking, selective attention, sustained attention, inhibition, focus-switching and updating may be the core executive processes involved in working memory. However, the exact involvement of these components depends on the measure of working memory in question. For instance, complex-span tasks (measuring working memory capacity) require the participant to remember stimulus sequences during an ongoing secondary task. This requires active maintenance of information in the face of concurrent processing demands and inhibition of task-irrelevant information. N-Back tasks use stimulus sequences such as letters, numbers or pictures and for each item in the

sequence participants judge whether the item matches with the N items back. This task requires the participants to actively maintain and update the information with each response as well as to inhibit the irrelevant information. Thus, an updating component is included in N-Back tasks and not in complex-span tasks.

The executive processes discussed so far are not specific to working memory tasks only, but are involved in other cognitive functions as well, including fluid intelligence. Miyake et al. (2000) identified task shifting, inhibition and updating as three related yet independent executive functions. Friedman et al. (2006) found that out of these three, only updating is correlated ($r = .74$, at the latent level) with measures of fluid intelligence such as Raven's Progressive Matrices Test (RAVENS; Raven, 1960) and the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1997) Block Design subtest. Their explanation is that updating requires an individual to sustain attention in order to process relevant and ignore irrelevant information, and in this manner, corresponds to Binet's definition of intelligence, which requires an individual to first perceive information, store the perception in memory and then retrieve it for further processing. Measures of fluid intelligence, such as RAVENS and Cattell's Culture Fair Test (CATTELL; Cattell, 1973) rely on the same ability to maintain activation of goal-relevant information and inhibition of distractors in addition to concurrent processing of information (Carpenter, Just, & Shell, 1990).

The idea behind recent working memory training is to improve the efficacy of these fundamental executive processes, with the hope that this will have an impact on other constructs that also employ some of the same underlying processes, specifically fluid intelligence. Since, updating is the overlapping executive process between working

memory and fluid intelligence, it is plausible that it is the crucial process that mediates the training benefits seen on measures of fluid intelligence. The current study specifically looked at the role of training the updating component in eliciting transfer effects to fluid intelligence and other executive components as defined by Miyake et al. (2000).

1.2 Process-Based Training

As mentioned earlier, recent working memory training studies involve an implicit, process-based approach where improvement in performance is based on task repetition and oftentimes on gradual adjustment of difficulty level (Klingberg, 2010). The idea here is to train the cognitive mechanisms underlying the training task, rather than train explicit strategies to improve performance. The results so far have demonstrated that process-based working memory training results in performance improvement on the trained tasks as well as some transfer of training on untrained tasks. The reported transfer effects have a very wide range: from reading (e.g. Chein & Morrison, 2010) to fluid reasoning (e.g. Jaeggi et al., 2008) to ADHD symptomatology (Beck, Hanson, Puffenberger, Benninger & Benninger, 2010) and to drinking behavior (Houben, Nederkoorn, Wiers, & Jansen, 2011). However, not all research groups have found significant transfer effects. That is, some studies have found that training only improves performance on the trained tasks and, to a lesser extent, on tasks that tap into the same construct as the trained tasks, but not to other cognitive constructs (e.g. Shipstead, Redick, & Engle, 2012). The similarities and differences in these studies, and the arguments for finding or not finding transfer effects are discussed below.

In a process-based training, participants are usually trained 4-5 days a week for 2-6 weeks on working memory tasks. Most of the training programs are also adaptive in

the difficulty of the tasks. That is, the training difficulty is increased when the participants perform well on the training task and decreased when performance is poor. This ensures an optimal level of performance from the participants. In addition, maintaining a difficulty level that matches the participant's developing skill levels to the demands of the task keeps the participant motivated to perform well. The tasks used for training vary in their complexity, from simple STM tasks such as repeating sequences in the correct order (e.g., Colom et al., 2010), to more complex working memory tasks such as identifying targets in an N-Back task in one or two modalities (e.g., Jaeggi et al., 2008), and training in working memory components such as switching between tasks (e.g., Karbach & Kray, 2009). It is expected that training-induced plasticity will affect a general mechanism or set of mechanisms that at least partially underlie both training and the transfer tasks.

The benefit of any cognitive training program can be assessed by gains on the trained tasks, transfer of training to untrained tasks, as well as the stability of training and transfer over time (Hertzog, Kramer, Wilson, & Lindenberger, 2008). The performance of the participants is assessed at least at two time points, once before the training and once after the training. The post-training performance on the trained tasks is compared to the pre-training performance to get an assessment of the training gains. Performance is also gauged on untrained tasks to assess transfer effects. These transfer effects can be either near transfer – to tasks measuring the same construct—or far transfer – to tasks measuring related but different constructs. More researchers have found support for near-transfer training effects than for far-transfer effects (e.g., Li et al., 2008).

Additionally, benefits can be gauged by comparing the post-training performance of the trained group against the performance of an active or passive control group. In an active control condition, participants are trained on a task not related to working memory; in a passive control condition the participants are only assessed at two time points, without any intervening cognitive training. These control groups help to assess if the effects of training seen in the trained group are due to motivational factors, experimenter bias, or any other non-specific factors.

In some studies, the maintenance of the training and transfer effect is examined by assessing the performance of the participants at a third time point, much later after training.

1.2.1 Process-Based Working Memory Training in Young Adults

The first few articles in the field of process-based cognitive training were by Klingberg and colleagues (Klingberg et al., 2005; Klingberg, Forssberg & Westerberg, 2002). Their studies focused on training children with ADHD on working memory tasks. The idea was to improve the ADHD symptoms and also obtain transfer on untrained working memory tasks as well as on tasks of fluid intelligence. The training resulted in performance gain on the trained tasks as well as transfer to untrained working memory tasks and tasks measuring fluid intelligence. In addition, ADHD symptoms also improved. Motivated by these findings, other researchers examined the effect of working memory training on participants of various age groups with or without cognitive deficits. The results of these studies were mixed in terms of finding transfer to untrained working memory tasks as well as to measures of fluid intelligence (e.g. Harrison et al., 2013; Jaeggi et al., 2008; Zinke et al., 2014).

Focusing on research with younger adults as training subjects, one of the seminal studies is by Jaeggi et al. (2008). The study consisted of four experiments in which the participants were trained for 25 minutes per session on adaptive dual N-Back tasks (visual and auditory). The four experiments differed in the number of sessions the participants were trained for (8, 12, 17 and 19 sessions). The results showed a dose-dependent training effect. That is, the more sessions the participants were trained for, the higher the training gains and transfer to fluid intelligence. Moreover, the gains were significantly higher for the trained groups compared to passive control groups. The authors attributed the transfer effects to the dual and adaptive nature of the training tasks. That is, the dual N-Back task engages multiple executive processes, including inhibiting irrelevant items, monitoring ongoing performance, task shifting, updating and binding; all of these processes are or might be correlated with measures of fluid intelligence. In addition, the adaptive nature of the training may have led to a continual engagement of these underlying executive processes in the N-Back task, minimizing the development of automatic processes and task-specific strategies. In a follow-up study, Jaeggi et al. (2010) explored if using a single N-Back task in place of the dual N-Back task would yield similar transfer effects. They found that both single and dual N-Back training groups showed transfer effects on matrix reasoning tests (RAPM and BOMAT) compared to a passive control group.

Several researchers have tried to replicate these findings, but not all have found significant transfer effects. For instance, Redick et al. (2013) replicated the dual N-Back study, adding an active control group to control for motivational factors, and including multiple measures of fluid intelligence, crystallized intelligence, perceptual speed,

multitasking, and working memory capacity. The performance on trained tasks improved for both trained and active control groups. However, no transfer was seen on any of the other cognitive abilities. In contrast, Jaeggi, Buschkuhl, Shah and Jonides (2014) did obtain transfer effect on a composite score of visuospatial measures as well as on BOMAT after incorporating Redick et al.'s (2013) changes in the original design. Importantly, they also found that subjects who believed in the malleability of intelligence showed greater transfer than those who did not. Thus, intrinsic motivation might be an important factor to obtain transfer effects in these kinds of training regimens.

Chein and Morrison (2010) compared participants trained on an adaptive complex working memory span task with a passive control group, and obtained mixed findings for transfer effects. Improved performance was seen on the working memory tasks, and transfer effects were obtained on Stroop task and reading comprehension. However, no performance increment was seen on fluid intelligence tasks and reasoning tasks. Heinzl et al. (2014) also found mixed results for transfer effects. In their study participants were trained on an adaptive N-Back task and their performance was contrasted against that of a passive control group. They found significant improvement on the training task and age-specific transfer effects. In young adults, transfer to executive functioning and processing speed was seen, whereas, in older adults, transfer was seen on STM, episodic memory and processing speed. Additionally, training did not transfer to fluid intelligence.

In contrast, Thompson et al. (2013) failed to find any near or far transfer effects after training participants on adaptive dual N-Back task, despite robust gains on the trained tasks. Chooi and Thompson (2012) too obtained a dose-response relationship after training on an adaptive dual N-Back task, but no evidence for transfer.

The studies discussed so far reflect the current state of research in this area. Though reliable training gains have been observed in all of these studies, transfer effects are still debatable. Meta-analyses conducted on working memory training literature have also generated mixed findings.

1.3 Meta-Analyses on Working Memory Training

Three meta-analyses (Au et al., 2014; Karbach & Verhaeghen, 2014; Melby-Lervag & Hulme, 2013) have looked at the effects of process-based working memory training in healthy younger adults.

Melby-Lervag and Hulme's (2013) analysis was based on twenty-three studies. The participants in these 23 studies were typically and atypically developing and developed individuals, ranging in age from ten to seventy-five. The working memory training interventions were all adaptive and administered for at least two weeks; all studies compared the treatment group with a control group. The meta-analysis found that the working memory training led to transfer on related working memory tasks (verbal working memory: $d = 0.78$; visuospatial working memory: $d = 0.52$). However, these performance improvements were short-lived and not seen when the participants were retested at a later point in time. Also, the benefit of training did not extend to tasks that were not related to working memory (verbal ability: $d = 0.13$; non-verbal ability: $d = 0.19$; Stroop task: $d = 0.32$; word decoding: $d = 0.13$; arithmetic: $d = 0.07$).

In comparison, Au et al. (2014) and Karbach and Verhaeghen (2014) found more promising results regarding the transfer effects of working memory training. Au et al.'s (2014) meta-analysis was based on twenty studies, which included only healthy participants between eighteen and fifty years of age. Only controlled studies based on

adaptive N-Back tasks and with fluid intelligence transfer measures were included. Their analysis found a small but significant effect of N-Back training on fluid intelligence ($g = 0.24$). Karbach and Verhaeghen (2014) included forty-nine studies in their analysis with healthy subjects in the age between 63 to 87 years of age; a subset of these studies (twenty-eight of them) also included younger adults (mean age 17 to 31). The participants were trained on working memory and executive functions (EF) tasks and transfer was assessed on tasks related and unrelated to these two cognitive constructs. The meta-analysis showed improved performance in both young and old adults on the trained tasks (young: $d = 0.78$; old: $d = 0.91$) as well as on the transfer tasks-- near (young: $d = 0.98$; old: $d = 0.47$) and far (young: $d = 0.72$; old: $d = 0.37$) compared to both active (young: $d = 0.34$; old: $d = 0.34$) and passive control groups (young: $d = 0.29$; old: $d = 0.09$).

1.4 Understanding the Underlying Mechanism in Working Memory Training:

Present study

N-Back is one of the most widely used working memory training tasks. Some studies using it in their training protocol have found transfer effects on the measures of fluid intelligence, others haven't. This study explored whether the differences in findings are due to the differential effect of training on specific executive-control components of the N-Back task, and the correlation of training-related changes with transfer tasks. Specifically, if the updating component is the key mediator in eliciting training and transfer effects in working memory training. Lilienthal, Tamez, Shelton, Myerson and Hale (2013) have explored a similar question. They examined which working memory component (in their definition: executive attention, updating, focus switching, increase in the capacity of the focus of attention, and short term memory) drives the training related

gains after dual N-Back training. They trained participants on an adaptive dual N-Back task and gave them a pre and post-training battery measuring each executive component of working memory. Based on the gain scores (post-pre) on transfer measures they suggested that an increase in the capacity of the focus of attention underlies the benefits of adaptive dual N-Back training. In Lilienthal et al.'s (2013) study, however, training-related gains on the specified components of the N-Back task were not calculated based on the training task. A more comprehensive exploration of this idea has been performed in this study by decomposing a modified version of the N-Back task (Price, Colflesh, Cerella, & Verhaeghen, 2014) and examining training-related changes on those decomposed components themselves. In addition, how the change in each component process correlates with the transfer effects on a variety of cognitive constructs was also studied. This study thus looked at how the component processes of the N-Back task change as a function of training and explain transfer effects in terms of those changes.

As stated earlier, performance of the participants undergoing the training protocol is usually compared against the performance of an active or passive control group to assess if the effects of training seen in the trained group are due to factors unrelated to the training program (including test-retest effects, motivational factors, experimenter bias, etc.). This study included an experimental group and an active control group designed to tease apart the role of updating in working memory training effects.

1.4.1 N-Back Task Decomposition

Price et al. (2014) performed a task decomposition of the modified N-Back task (Verhaeghen, Cerella, & Basak, 2004) and showed that each of its executive components (listed below) can be evaluated separately. In their modified N-Back task – which was

used here as well -- items appear from left to right in N columns and subjects respond whether the item shown in the column matches to the item previously shown in the same column. Price et al. used Cowan's (2005) embedded-process model of working memory comprising of an inner store (capacity-limited area of immediate access) and outer store (activated portion of long-term memory with no capacity limitation) to decompose the N-Back task. Information in the outer store is not directly accessible and has to be accessed by transferring it into the inner store (also called focus of attention). When the capacity of inner store is reached, items will be removed to the outer store to allow for processing of a new item. This executive component of working memory is called updating. Similarly, an item may be required to be brought back to focus of attention from outer store for processing. This process is called focus switching. If the item retrieval order is the same as the study order, focus switching is a direct one-step retrieval process (Verhaeghen et al., 2004). However, if the item retrieval order is reverse or random compared to the study order it requires an additional search process in the outer store (Lange, Cerella, & Verhaeghen, 2011). Thus, the updating process, the focus-switching one-step retrieval process, and the focus switching requiring search are the three executive components outlined by Price et al. in the N-Back task. Price et al (2014) trained participants for 10 sessions on the modified N-Back task and found training-related improvements on the updating process, the direct retrieval process, as well as an increase in the capacity of the focus of attention (for the latter result, see also Verhaeghen et al., 2004), but not on the search process.

1.4.2 Study Details

To measure the three executive components -- updating, focus switching one-step retrieval and focus switching requiring search, as well as on the size of the focus of attention, three versions of the modified N-Back task (Figure 1) were used. In the first version (forward-updating) items appeared from left to right in N columns and subjects responded whether the item shown in the column matched the item previously shown in the same column. In the second version (forward-no-updating) participants matched all the probes with the first set of items shown in the respective N column. The third version (random-no-updating) was similar to the second; however, the participants were probed on N items in random order. In the random-no-updating task, a trial was classified as a non-switch trial if the probe location was the same as the immediately preceding probe location; a switch trial occurred when the probe location was different from the immediately preceding probe location.

Component scores were calculated as in Price et al. (2014). The duration of the one-step-direct retrieval operation was estimated from the forward no-updating condition as the increment in RT from $N=1$ to $N=2$. Duration of the random search process was estimated by the difference between the focus-switch and non-switch probes in the random-no-updating condition. The duration of the updating process was estimated as the difference between the forward-updating and forward-no-updating conditions. To study the effects of training on the size of the focus of attention, each level of N was evaluated as a function of training in the forward-no-updating condition. Each level of N was compared to the next higher level of N (one comparison set) to see if the increase in RT was significantly different. If the RT difference in a comparison set became significant, the lower level of N signified the size of the focus of attention.



Figure 1 - Example 3-Back tasks: FWNU -- forward access without updating; FWUP -- forward access with updating; RNNU -- random access without updating (Price et al., 2014)

Participants in both the experimental and active control groups were trained for about 10 hrs. over a period of two weeks (5 days a week) in the lab. Experimental group was trained on all the three versions of the task, whereas, the active control group was trained on the forward-no-updating and random-no-updating tasks. Thus, the active control group was trained on all the sub-components of the modified N-Back task, except updating. The effect of the training was assessed on the three executive components - updating, focus switching one-step retrieval and focus switching requiring search -- as well as on the size of the focus of attention. Correlation between the training related changes in the executive components (including the size of the focus of attention) and the transfer measures was also evaluated.

1.4.2.1 Transfer Measures

Transfer effects were assessed for the constructs of updating, task switching, inhibition, focus switching and fluid intelligence. These constructs were selected based on either being part of the component processes of working memory or being used most commonly in the working memory training literature to assess transfer. Each construct was measured by multiple tasks to obtain a reliable and valid estimate of the construct. For updating, task switching and inhibition, tasks were selected based on Miyake et al. (2000). For a description of the actual tasks used, see the Methods section. In the pre-training session, participants completed the demographic questionnaire and all the pre-training assessment on the transfer measures. Participants took about two hours to complete this session. They started the training from the next day onwards. After completing all 10 training sessions, participants came back the next day for the post-training session. The order of the tasks in pre and post training sessions was the same. All the tasks used parallel forms or split items for the pre-post assessments.

In addition, an open-ended strategy questionnaire was administered after training sessions 1, 3, 5, 7, 9 and 10 to examine if participants developed strategies to perform better on the tasks. Jaeggi et al., (2014) found that the training and transfer effects seen in some studies may be due to the belief of the participants about malleability of cognition. If participants believe that they can train their cognition, this may lead to higher transfer scores. To examine this hypothesis, participant filled out the Theories of Cognitive Abilities scale (TOCA, Dweck, 1999).

CHAPTER 2. METHOD

2.1 Participants

Sixty-two participants between the ages of 18 to 30 years were initially recruited from the Georgia Institute of Technology. Participants were compensated \$155 for completing the study. Data from four participants were excluded from the analysis since they did not complete the study. Thus, the final sample consisted of 29 participants (age: $M = 20.26$, $SD = 2.68$; education in years: $M = 13.9$, $SD = 1.63$; 14 women) in the experimental group and 29 participants (age: $M = 20.48$, $SD = 1.86$; education in years: $M = 14.28$, $SD = 1.58$; 9 women) in the active control group. The groups were not significantly different in age, $t(56) = 0.34$, $p > 0.05$ and education, $t(56) = 0.9$, $p > 0.05$. All participants gave written informed consent before they started the experiment.

2.2 Training Task

For each training session all three versions of the modified N-Back task (Figure 1) were given to the participants in the experimental group, whereas only forward-no-updating and random-no-updating versions were given to the active control group. Presentation of these versions was blocked, and their order was counterbalanced between participants. At the beginning of each trial, participants were required to encode N digits shown on the screen in separate columns, and hit the spacebar when they were ready to begin the trial. Subsequent digits appeared on the screen one at a time and participants made a judgment about whether the current digit matched a digit previously shown in the same column. They pressed the “1” key if it was a match and the “3” key if it was a

mismatch. In the *forward-no-updating* task participants compared the digit shown on the screen to the digit originally encoded in the same column. In the *forward-updating* task participants compared the digit to the digit most recently presented in the same column. In both these conditions probe digits appeared left-to-right in a sequential format. The *random-no-updating* task was identical to the forward-no-updating condition, except that the probe digits randomly appeared in any column, and thus not followed the sequential format. In this task, the digit probe appeared in the same location as the previous digit in 50% of the trials (non-switch trials).

For each level of N there were 31 trials constituting 1 block, except for N=1, for which there were 11 trials; training for each training task consisted of 10 blocks. The first trial of each block was considered a practice trial and not included in the analysis. For the first session, training began at the N = 1 level; for subsequent sessions, N was set at 2 values below the highest level of N reached in the previous session. After completing one block, participants were given feedback on their accuracy and RT. If their accuracy was at 90% or higher, difficulty was increased by one value of N; if their accuracy was at 70% or lower, difficulty was decreased by one value of N (Jaeggi et al., 2008). The highest N-Back level that a participant could reach was N = 13, due to screen size limitations. Across a block, match/mismatch probes were equally likely to occur. After completing 10 blocks, participants moved on to the next training task. In the first session, participants were given 10 practice trials for each training task, at N = 3. For the random-no-updating task, N = 1 was omitted, as this would be identical to the forward-no-updating task. Participants completed 10 training sessions lasting approximately an hour each over a period of two weeks. Accuracy and RT was recorded. At the end of session 1,

3, 5, 7, 9 and 10 participants were asked if they were using any specific strategy to complete the task.

2.3 Pre-Post Assessments

To study the effect of training on target and transfer measures, participants were assessed pre and post training on the modified N-Back task as well as measures of updating, task switching, inhibition, focus switching and fluid intelligence. In addition, the TOCA (Dweck, 1999) scale was also administered.

2.3.1 Modified N-Back Task

Participants in both groups performed the modified N-Back task (described above) with N ranging from 1-5 (31 trials per block for N = 2-5, 11 trials per block for N = 1; 2 blocks for each N) for each task condition (forward-no-updating, random-no-updating and forward-updating).

2.3.2 Updating

2.3.2.1 Keep-Track Task (Miyake et al., 2000; Yntema, 1963)

In this task, participants were shown 16 words sequentially from six semantic categories (animals, colors, countries, distances, metals, and relatives) and asked to remember the last word presented for each of a number of target categories. For example, if the target categories were metals, relatives, and countries, then, at the end of the trial, participants recalled the last metal, the last relative, and the last country presented in the list. Participants were given three trials with three target categories and three with four

target categories. Before each trial, subjects were told the categories that they should identify and recall the most recent instances from. The dependent variable was the number of words recalled correctly (maximum score 21). Percentage accuracy was calculated for group analysis.

2.3.2.2 Letter Memory Task (Miyake et al., 2000; Morris & Jones, 1990)

In this task, participants were shown lists of letters sequentially and asked to recall the last four letters presented in each list. The number of letters in each list varied (5, 7, 9, or 11) randomly across trials. There were 12 trials and the dependent measure was the number of correctly recalled letters (maximum score 48). Percentage accuracy was calculated for group analysis.

2.3.3 *Task Switching*

2.3.3.1 Plus–Minus Task (Miyake et al., 2000; Spector & Beiderman, 1976)

During this task, participants were asked to complete problems from three different lists (consisting of 30 two-digit numbers each) as quickly and accurately as possible. In the first list, the subjects added three to each number; in the second list, the subjects subtracted three from each number; in the third list, the subjects alternated between adding three and subtracting three from each number. The first two lists did not require switching between the tasks, the third did. The difference between the RT of the third list and the average RT of first two lists was the switching cost. Median RT was calculated for each participant for group analysis.

2.3.3.2 Number-Letter Task (Miyake et al., 2000; Rogers & Monsell, 1995)

Participants were shown a number–letter pair (e.g., 5K) in one of the four quadrants on the computer screen. The task was to indicate whether the number was odd or even (2, 4, 6, and 8 for even; 3, 5, 7, and 9 for odd) when the number–letter pair was presented in either of the top two quadrants and whether the letter was a consonant or a vowel (G, K, M, and R for consonant; A, E, I, and U for vowel) when the number–letter pair was presented in either of the bottom two quadrants. Participants were given 100 trials; half of them required a switch between upper and lower quadrants and were tagged as switch trials. The order of trials was random. The switching cost for this task was the difference between the RT of the switch and non-switch trials. Median RT was calculated for each participant for group analysis.

2.3.4 Inhibition

2.3.4.1 Stroop Task (Stroop, 1935)

In this task, participants were asked to identify the color of the stimulus as quickly as possible and respond by pressing the key representing that color. The task included 60 trials with a string of asterisks printed in one of six colors (red, green, blue, orange, yellow, or purple) and 60 trials with a color word printed in a different color (e.g., BLUE printed in red color). Presentation of these different trial types was blocked. The dependent measure was the RT difference between the trials in which the word and the color were incongruent and the trials that consisted of asterisks. Median RT was calculated for each participant for group analysis.

2.3.4.2 Antisaccade task (Miyake et al., 2000; Roberts, Hager, & Heron, 1994)

In this task, participants saw a fixation cross at the center of the screen for a variable amount of time (between 1,500 and 3,500 ms, in 250-ms intervals) followed by a distractor (blank square) on one side of the screen for 225 ms and then the target stimulus (arrow inside an open square) on the opposite side of the screen for 150 ms. The participant's task was to indicate the direction of the arrow (left, up, down, or right). There were 90 target trials and the number of correct responses was the dependent measure. Percentage accuracy was calculated for group analysis.

2.3.4.3 Stop-signal task (Logan, 1994)

The task consisted of two blocks of trials. The first block was used to build up a prepotent categorization response. Participants were presented with 24 monosyllabic words (matched for length and frequency) sequentially at the center of the screen for 1,000 ms each. Their task was to verbally categorize each word as an animal or a non-animal. They were given 2,000 ms to do so. In the second block of 48 trials the procedure was the same, except that the participants were required to not respond (inhibit categorization) when three asterisks were presented below the word. Asterisks were presented on 16 of the trials. Participants were instructed not to slow down to wait for possible signals. The score was the number of categorization responses given to the "stop" trials. Percentage accuracy was calculated for group analysis.

2.3.5 *Focus Switching*

2.3.5.1 Garavan Task (Garavan, 1998)

In this task, participants saw a series of triangles and rectangles, presented one at a time at the center of screen. The rectangles could be in the horizontal or the vertical orientation and the triangle could be pointed downwards or upwards. The length of the sequences varied between 16-19 stimuli. There were 2 trials of each sequence length, resulting in a total of 8 trials. The order of the stimuli within a sequence was random. The task was to keep separate running counts of how many triangles and rectangles were presented in each sequence. RT was measured from the presentation of a shape to when the participant pressed the space bar to see the next shape. Switch trial RTs (rectangle followed by a triangle or vice-versa) minus repeat trial RTs (repetition of the same shape stimulus) for the correct trials was used as a measure of focus switching. Median RT was calculated for each participant for group analysis.

2.3.5.2 Spatial Focus Switching (Oberauer, 2003)

In this task, participants were presented with two frames on the screen, each with a digit. They studied the digits and then hit the spacebar to continue. Immediately afterwards, a mathematical operation was displayed in one of the frames. Operations ranged from “-8” to “+8” (excluding 0), with the restriction that the result will be a digit between one and nine. Participants were asked to respond with the correct answer as quickly as possible. Each response was immediately followed by another operation displayed in the same or another frame. There were 9 operation tasks in each trial, half of which were switch trials. Participants completed 16 trials. Switch cost (RT switch trials- RT non-switch trials) was the dependent variable. Median RT was calculated for each participant for group analysis.

2.3.6 *Fluid Intelligence*

2.3.6.1 RAPM (Raven, 1990)

Participants were shown a 3 x 3 matrix of abstract shapes and patterns with the shape in the bottom right location missing. Their task was to select an item from eight possible choices that best completes the overall pattern. The test, which consists of 36 items, was divided into two parts (Jaeggi et al., 2014). One set was administered at pre-test and other at post-test for each participant. The order of administration of the version of the test was counterbalanced across participants. Subjects had 10 min to complete 18 questions. The number of correct responses out of 18 was used as the dependent variable. Percentage accuracy was calculated for group analysis.

2.3.6.2 BOMAT (Hossiep, Turck, & Hasella, 1999)

In this task participants were shown a 5 x 3 matrix of abstract shapes with one of the slots empty. Their task was to select an item from six possible choices that best completes the overall pattern. Form A from the short version was used. The items in the test were divided into even and odd questions. One set was administered at pre-test and other at post-test for each participant. Participants completed 15 questions in 12 min. The dependent variable was the number of correctly solved problems. Percentage accuracy was calculated for group analysis.

2.3.7 *Theories of Cognitive Abilities (TOCA; Dweck, 1999)*

The TOCA assessed the degree to which participants believe that intelligence is malleable. It consists of eight items, scored on a six-point Likert scale. It was administered only at the pre training assessment.

CHAPTER 3. RESULTS

3.1 Training Tasks

To calculate the gains on the trained tasks, the last level of N reached for each training session was used. Overall, the experimental and the active control group showed an increase in their performance on the trained tasks (Table 1).

For each training task, a three-parameter negative exponential function ($y = a - b \cdot \exp(-c \cdot \text{session})$) was fitted to the averaged data separately for each task and group (Gashler, Progscha, Smallbone, Ram, & Bilalic, 2014) (Figure 2). The three parameters (Table 2) used in the fitting function are the asymptote, or final level of performance (a), the difference between initial and final performance, or the amount of learning (b), and curvature, or learning slope (c). Compared to the forward-updating task, the amount of learning (b) was much larger, the learning slope (c) mostly steeper (random no-updating in the experimental group was the exception), and the asymptote (a) much higher for the forward-no-updating and the random-no-updating tasks. Thus, the forward-updating task was more difficult to master on all three parameters of the learning curve.

Table 1 - Last N level reached after each training session for Forward-no-updating (FWNU), Random-no-updating (RNU) and Forward-updating (FWU) conditions in the experimental group (exp) and the active control group (active).

	Session									
	1	2	3	4	5	6	7	8	9	10
FWNU_exp	9.38	11.69	11.90	12.17	12.10	12.28	12.21	12.41	12.34	12.41
FWNU_active	9.24	11.93	11.69	12.48	12.41	12.48	12.69	12.83	12.83	12.83

Table 1 (continued)

RNU_exp	9.72	11.07	11.52	11.72	11.76	12.00	11.86	12.17	12.31	12.38
RNU_active	9.90	12.14	12.48	12.52	12.59	12.69	12.69	12.83	12.66	12.76
FWU_exp	6.41	7.21	7.14	7.48	7.62	7.59	7.59	7.90	8.17	8.17

Table 2 - Fit parameters after fitting negative exponential function ($y = a-b*\exp(-c*x)$) to the training data of Forward-no-updating (FWNU), Random-no-updating (RNU) and Forward-updating (FWU) conditions in experimental (exp) and active control groups (active)

		<i>Rsquare</i>	<i>a (95% CI)</i>	<i>b (95% CI)</i>	<i>c (95% CI)</i>
FWNU	exp	0.98	12.27 (12.14, 12.4)	11.86 (5.94, 17.79)	1.42 (0.94, 1.9)
	active	0.94	12.65 (12.35, 12.95)	9.56 (3.21, 15.92)	1.06 (0.45, 1.66)
RNU	exp	0.96	12.17 (11.94, 12.4)	4.33 (2.87, 5.79)	0.6 (0.34, 0.88)
	active	0.99	12.68 (12.59, 12.76)	13.08 (7.8, 18.36)	1.55 (1.16, 1.94)
FWU	exp	0.90	8.42 (7.26, 9.58)	2.2 (1.37, 3.03)	0.19 (-0.06, 0.44)

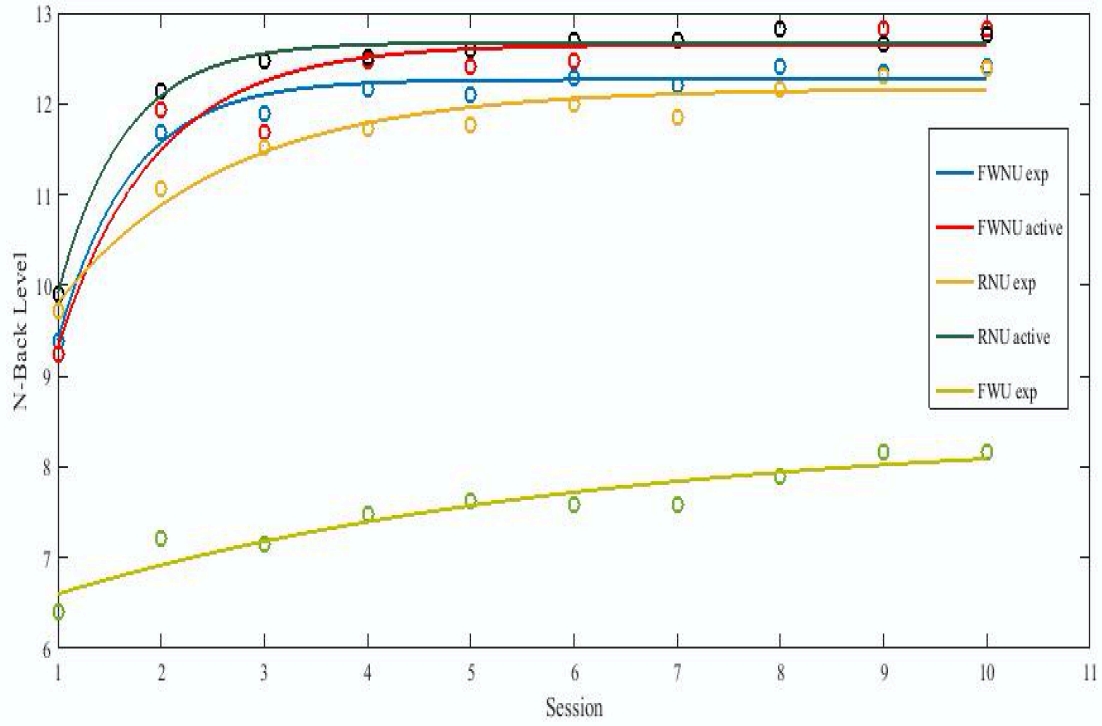


Figure 2 - Negative exponential function fitted to the training data of Forward-no-updating (FWNU), Random-no-updating (RNU) and Forward-updating (FWU) conditions in the experimental (exp) and active control groups (active). The circles represent the actual N-Back level reached after each training session. Solid lines represent the fitted curve.

3.2 Pre and Post Training Assessment of the N-Back Component Processes

3.2.1 Trained Tasks

Means RTs and standard deviations (SD) for the three N-Back tasks and the different N-Back components were calculated at pre and post training time points. Accuracy for each task was 95% or higher and hence not analyzed. To assess the effect of training on different levels of N and different training groups 2 (group: experimental and active control) x 2 (time: pre and post) x 5 (N:1-5) repeated measures ANOVA was

conducted on the forward-no-updating and the forward-updating tasks; 2 (group: experimental and active control) x 2 (time: pre and post) x 4 (N:2-5) repeated-measures ANOVA was conducted separately on the switch and non-switch trials of the random-no-updating task.

Compared to the pre training scores, both training groups showed a marked improvement (lower RTs) on all the tasks that they were trained on. For the forward-no-updating task (Figure 3) we found a significant main effect of time, $F(1, 56) = 281.36, p < 0.05, \eta^2p = 0.83$. Compared to pre training, ($M = 731.69, SD = 142.6$), the RTs were much lower at post-training assessment, ($M = 471.69, SD = 69.92$). There was also a main effect of level of N, $F(4, 224) = 12.39, p < 0.05, \eta^2p = 0.18$. Comparison was made between RTs of 1-Back with 2-Back, 2-Back with 3-Back, 3-Back with 4-Back and 4-Back with 5-Back. After Bonferroni correction, we found that RT for 2-Back level ($M = 617.85, SD = 99.76$) was significantly higher compared to 1-Back ($M = 560.15, SD = 121.46$), $t(57) = 4.76, p < 0.05$. Interestingly, RT for 2-Back was also significantly higher compared to 3-Back ($M = 602.77, SD = 98.67$), $t(57) = 2.68, p < 0.05$. There was no significant difference between 3-Back and 4-Back ($M = 612.41, SD = 102.29$), $t(57) = 1.43, p = 0.16$, or between 4-Back and 5-Back ($M = 615.27, SD = 107.58$), $t(57) = 0.55, p = 0.58$. In addition, there was no main effect of group, $F(1, 56) = 0.07, p = 0.8, \eta^2p = 0.001$, i.e. RTs for both the active control ($M = 604.99, SD = 109.59$) and the experimental groups ($M = 598.39, SD = 81.47$) were similar while performing the task. There were no significant interactions, all F values < 2.4 , largest $\eta^2p = 0.04$. Thus, both groups benefited equally from the forward-no-updating training.

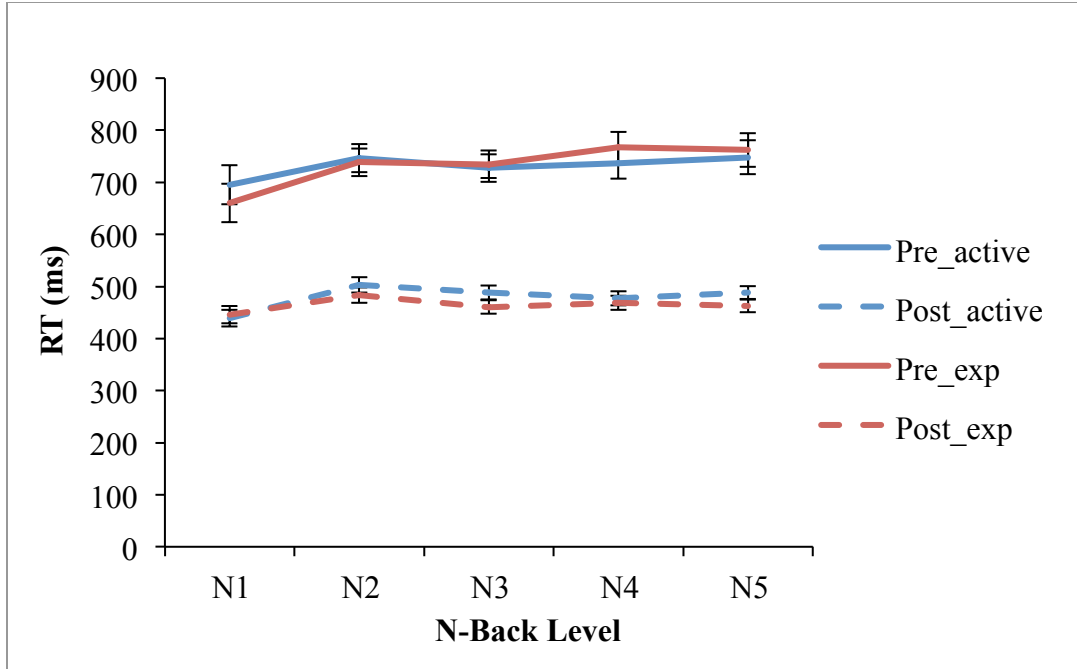


Figure 3- Figure shows the RT data for the forward-no-updating task at pre and post training assessment for the active control (active) and the experimental (exp) group. There was a significant main effect of time and N-Back level. Error bars denote the standard error of the mean.

For the non-switch trials of the random-no-updating task (Figure 4) we found main effects of time, $F(1, 56) = 144.38, p < 0.05, \eta^2p = 0.72$, and N-Back level, $F(3, 168) = 20.13, p < 0.05, \eta^2p = 0.26$. Participants were significantly faster at the task at post training assessment ($M = 500.02, SD = 185.54$) compared to pre training assessment ($M = 724.54, SD = 90.18$). We compared RTs of 2-Back with 3-Back, 3-Back with 4-Back and 4-Back with 5-Back. After Bonferroni correction, we found significant differences between 3-Back ($M = 599.23, SD = 117.12$) and 4-Back ($M = 620.33, SD = 142.54$), $t(57) = 3.62, p < 0.05$, as well as between 4-Back and 5-Back ($M = 645.98, SD = 152.48$), $t(57) = 3.0, p < 0.05$. Thus, there was a linear increase in RT as the N-Back level

increased. We did not find main effect of group, $F(1, 56) = 0.24, p = 0.63, \eta^2_p = 0.004$. Thus, the performance of the active control ($M = 604.05, SD = 100.59$) and the experimental group ($M = 620.53, SD = 151.13$) was similar on this task. There was also a significant interaction between time and N-Back level, $F(3, 168) = 26.84, p < 0.05, \eta^2_p = 0.32$. There was a significant difference between 2-Back pre ($M = 662.26, SD = 162.77$) and post ($M = 504.94, SD = 104.48$), $t(57) = 8.48, p < 0.05$; 3-Back pre ($M = 700.76, SD = 171.36$) and post ($M = 497.71, SD = 85.86$), $t(57) = 11.76, p < 0.05$; 4-Back pre ($M = 739.75, SD = 213.61$) and post ($M = 500.91, SD = 96.00$), $t(57) = 10.79, p < 0.05$; 5-Back pre ($M = 795.41, SD = 234.96$) and post ($M = 496.56, SD = 100.14$), $t(57) = 11.76, p < 0.05$; 2-Back pre and 3-Back pre, $t(57) = 3.59, p < 0.05$; 3-Back pre and 4-Back pre, $t(57) = 3.74, p < 0.05$; 4-Back pre and 5-Back pre, $t(57) = 3.42, p < 0.05$. However, there was no difference between 2-Back post and 3-Back post, $t(57) = 0.76, p = 0.45$; 3-Back post and 4-Back post, $t(57) = 0.60, p = 0.55$; 4-Back post and 5-Back post, $t(57) = 0.88, p = 0.39$. Thus, with training the RT decreased for all the N-Back levels. There was a significant linear increase in RT with N at pre training assessment; however, this increase was not significant after training, suggesting the component reached perfect efficiency. None of the other interactions were significant, all F values < 0.72 , largest $\eta^2_p = 0.01$.

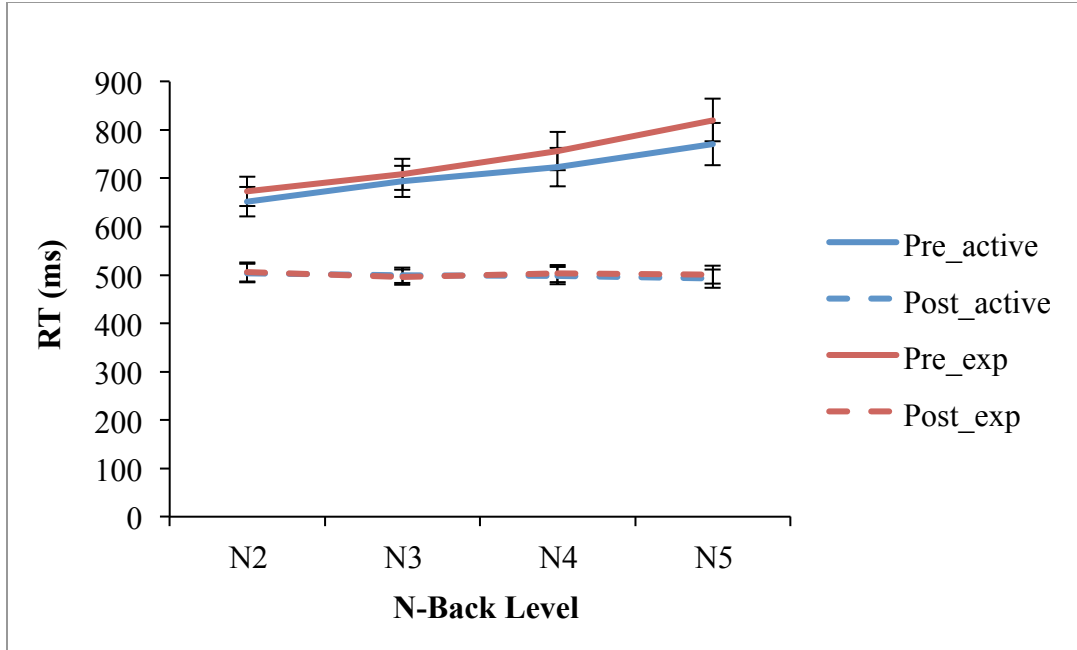


Figure 4 - Figure shows the RT data for the non-switch trials of the random-no-updating task at pre and post training assessment for the active control (active) and the experimental (exp) group. Main effect of time and N-Back level was significant. In addition, interaction between time and N-Back level was also significant. Error bars denote the standard error of the mean.

For the switch trials of the random-no-updating task (Figure 5) we found main effect of time $F(1, 56) = 81.41, p < 0.05, \eta^2p = 0.59$ and N-Back, $F(3, 168) = 315.26, p < 0.05, \eta^2p = 0.85$. Participants RTs were faster at post training ($M = 683.91, SD = 107.72$) assessment compared to the pre training ($M = 930.11, SD = 231.36$) assessment. We compared RTs of 2-Back with 3-Back, 3-Back with 4-Back and 4-Back with 5-Back. After Bonferroni correction, we found significant difference between 2-Back ($M = 639.75, SD = 127.86$) and 3-Back ($M = 726.07, SD = 143.12$), $t(57) = 9.94, p < 0.05$, 3-Back and 4-Back ($M = 844.66, SD = 168.56$), $t(57) = 12.46, p < 0.05$, as well as between 4-Back and 5-Back ($M = 1017.57, SD = 194.24$), $t(57) = 12.81, p < 0.05$. Thus, there was

again a linear increase in RT as the N-Back level increased. We did not find main effect of group, $F(1, 56) = 0.12, p = 0.73, \eta^2p = 0.002$. Thus, the performance of the active control ($M = 800.18, SD = 151.91$) and the experimental group ($M = 813.84, SD = 146.47$) was similar on this task. There was also a significant interaction between time and N-Back level, $F(3, 168) = 25.90, p < 0.05, \eta^2p = 0.32$. Post hoc analysis showed that with training RT decreased for all the N-Back levels and there was a significant linear increase in RT with N at pre and post training assessment. There was a significant difference between 2-Back pre ($M = 726.75, SD = 192.20$) and post ($M = 552.75, SD = 110.87$), $t(57) = 7.28, p < 0.05$; 3-Back pre ($M = 825.25, SD = 223.49$) and post ($M = 626.89, SD = 101.65$), $t(57) = 7.67, p < 0.05$; 4-Back pre ($M = 976.67, SD = 260.43$) and post ($M = 712.64, SD = 125.80$), $t(57) = 8.68, p < 0.05$; 5-Back pre ($M = 1191.77, SD = 301.73$) and post ($M = 843.36, SD = 165.63$), $t(57) = 9.05, p < 0.05$; 2-Back pre and 3-Back pre, $t(57) = 8.93, p < 0.05$; 3-Back pre and 4-Back pre, $t(57) = 9.68, p < 0.05$; 4-Back pre and 5-Back pre, $t(57) = 10.47, p < 0.05$; 2-Back post and 3-Back post, $t(57) = 5.91, p < 0.05$; 3-Back post and 4-Back post, $t(57) = 9.12, p < 0.05$; 4-Back post and 5-Back post, $t(57) = 8.44, p < 0.05$. Thus, the load effect remained the same even after training for both groups, suggesting that the switch trials are less trainable. N-Back level x group interaction was also significant, $F(3, 168) = 3.01, p < 0.05, \eta^2p = 0.05$. However, post-hoc analysis did not yield any theoretically relevant significant results. None of the other interactions were significant, all F values < 1.24 , largest $\eta^2p = 0.02$.

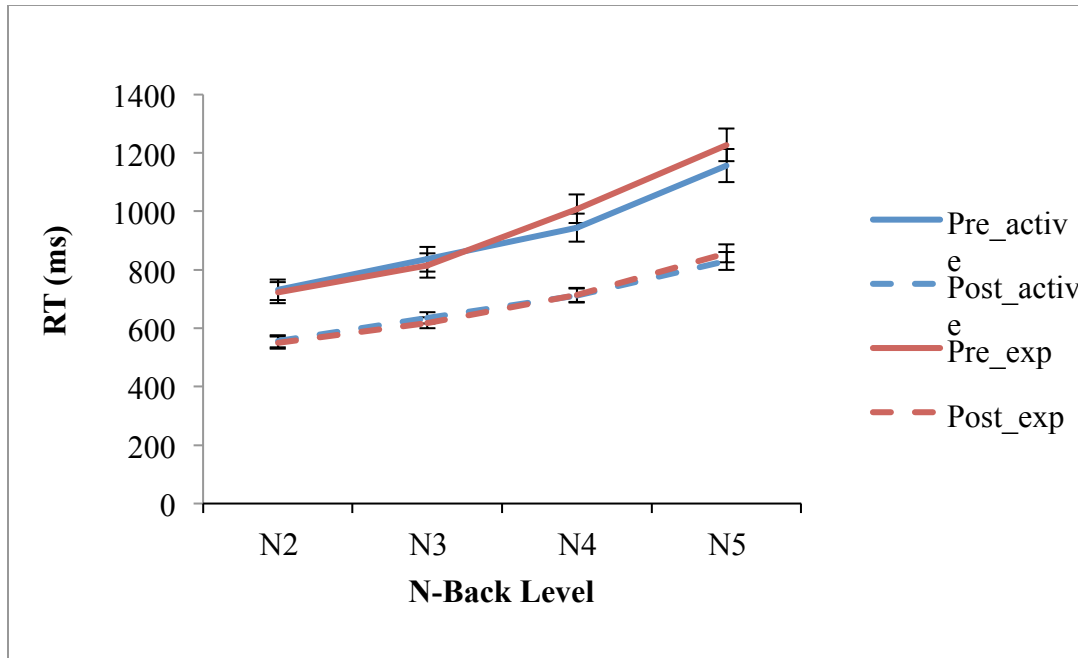


Figure 5 - Figure shows the RT data for the switch trials of the random-no-updating task at pre and post training assessment for the active control (active) and the experimental (exp) group. Significant effect of time, N-Back level and interaction between time and N-Back level was present. Error bars denote the standard error of the mean.

In the forward-updating task (Figure 6), there was a significant main effect of time, $F(1, 56) = 82.4, p < 0.05, \eta^2p = 0.60$, N-Back level, $F(4, 224) = 21.46, p < 0.05, \eta^2p = 0.28$ as well as group, $F(1, 56) = 4.06, p < 0.05, \eta^2p = 0.07$. Participants became faster with training on the updating task (pre: $M = 1224.84, SD = 385.90$; post: $M = 817.28, SD = 272.12$). To understand the main effect of N-Back level we compared RTs of 1-Back ($M = 881.60, SD = 236.58$) with 2-Back ($M = 970.41, SD = 241.60$), 2-Back with 3-Back ($M = 1018.82, SD = 256.93$), 3-Back with 4-Back ($M = 1092.79, SD = 373.89$) and 4-Back with 5-Back ($M = 1141.67, SD = 418.03$). After Bonferroni correction, we found significant difference between 1-Back and 2-Back, $t(57) = 3.97, p <$

0.05; 2-Back and 3-Back, $t(57) = 2.77$, $p < 0.05$; 3-Back and 4-Back, $t(57) = 2.75$, $p < 0.05$. Thus, there was a linear increase in RT with N. Note that the active control group ($M = 1092.02$, $SD = 290.60$) was not trained on the updating task, thus their RT was significantly higher than that of the experimental group ($M = 950.10$, $SD = 243.81$), $t(56) = 2.01$, $p < 0.05$. In addition, there was a significant interaction between time and group, $F(1, 56) = 13.57$, $p < 0.05$, $\eta^2_p = 0.56$. The two groups had similar RT at pre training (active control: $M = 1213.09$, $SD = 408.26$; experimental: $M = 1236.59$, $SD = 369.04$), $t(56) = 0.23$, $p = 0.82$, however post training the experimental group ($M = 663.62$, $SD = 166.26$) had significantly lower RT compared to the active control group ($M = 970.95$, $SD = 272.38$), $t(56) = 5.19$, $p < 0.05$. Thus, there was a differential effect of training on the updating task. There was also a significant interaction between time and N-Back level, $F(4, 224) = 5.84$, $p < 0.05$, $\eta^2_p = 0.10$. Bonferroni-corrected post hoc analysis showed significant difference between 1-Back pre ($M = 1018.00$, $SD = 327.68$) and 1-Back post ($M = 745.20$, $SD = 263.37$), $t(56) = 5.77$, $p < 0.05$; 2-Back pre ($M = 1153.68$, $SD = 362.91$) and 2-Back post ($M = 787.14$, $SD = 245.17$), $t(56) = 7.20$, $p < 0.05$; 3-Back pre ($M = 1230.72$, $SD = 362.77$) and 3-Back post ($M = 806.92$, $SD = 264.48$), $t(56) = 8.66$, $p < 0.05$; 4-Back pre ($M = 1236.76$, $SD = 520.21$) and 4-Back post ($M = 858.83$, $SD = 370.02$), $t(56) = 7.05$, $p < 0.05$; 5-Back pre ($M = 1395.03$, $SD = 624.95$) and 5-Back post ($M = 888.32$, $SD = 381.65$), $t(56) = 6.32$, $p < 0.05$; 1-Back pre and 2-Back pre, $t(56) = 3.51$, $p < 0.05$. Participants were significantly faster at post training assessment for all N-Back levels compared to the pre training assessment. In addition, even though there was a linear increase in RT as the N-Back level increased at pre training assessment, this increase was significantly more in going from 1-Back to 2-Back. Group x N-Back level

interaction was also significant, $F(4, 224) = 2.58, p < 0.05, \eta^2p = 0.04$. However, after Bonferroni correction none of the post hoc comparisons were significant. The three way interaction between N-Back level, group and time was also not statistically significant, $F(4, 224) = 1.23, p = 0.30, \eta^2p = 0.02$.

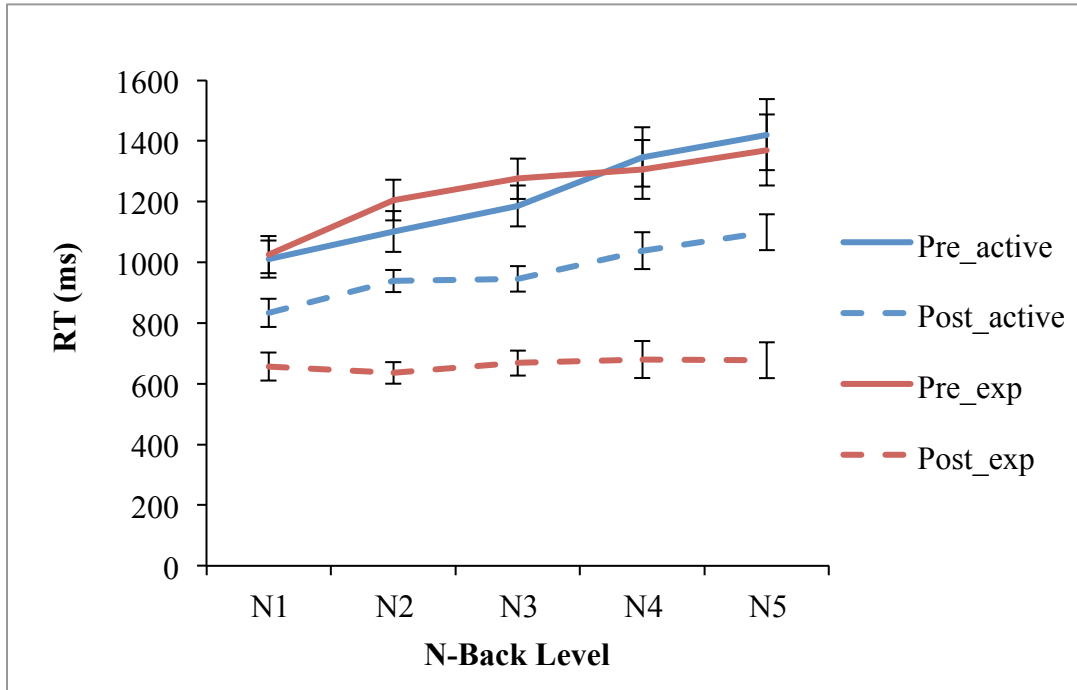


Figure 6 - Figure shows the RT data for the forward-updating task at pre and post training assessment for the active control (active) and the experimental (exp) group. There was a significant main effect of time, N-Back level and group, as well as significant interaction between time and group, and between time and N-Back level. Error bars denote the standard error of the mean.

3.2.1.1 N-Back Component Processes

As stated earlier, the modified N-Back task can be decomposed into focus switching direct retrieval process, focus switching requiring search process and the updating process. In addition, we can also calculate the size of focus of attention. Effect

of training was evaluated on these three N-Back component processes as well as on the size of focus of attention.

A 2 (time: pre and post) x 2 (group: active control and experimental) repeated-measures ANOVA was performed on the focus switching direct retrieval component (calculated as RT difference between 2-Back and 1-Back levels in the forward-no-updating task); 2 (time: pre and post) x 2 (group: active control and experimental) x 4 (N: 2-5) repeated-measures ANOVA was performed on the focus switching requiring search component (calculated as RT difference between switch and non-switch trials in the random-no-updating task); 2 (time: pre and post) x 2 (group: active control and experimental) x 5 (N: 1-5) repeated-measures ANOVA was done on the updating component (calculated as the difference in RT between forward-updating and the forward-no-updating task). For calculating the size of the focus of attention, first a 2 (time: pre and post) x 2 (group: active control and experimental) x 5 (N: 1-5) repeated measures ANOVA was done on the forward-no-updating data to see if there is any difference with respect to time, N and group. If there was a significant main effect of N and time, paired t-tests were to be performed comparing each level of N with the next higher level of N. This was done separately for the pre and post training forward-no-updating data.

For the focus switching direct retrieval process (Figure 7) we found no main effect of time, $F(1, 56) = 0.37, p = 0.54, \eta^2_p = 0.01$ or group, $F(1, 56) = 0.001, p = 0.99, \eta^2_p = 0.001$. The interaction between time and group was also not significant, $F(1, 56) = 1.40, p = 0.24, \eta^2_p = 0.02$. Thus there was no significant change in the focus switching direct retrieval process (calculated as RT difference between 2-Back and 1-Back levels in

the forward-no-updating task) as a function of training, suggesting a lack of malleability in the component. This pattern of result was same for the experimental and the active control group.

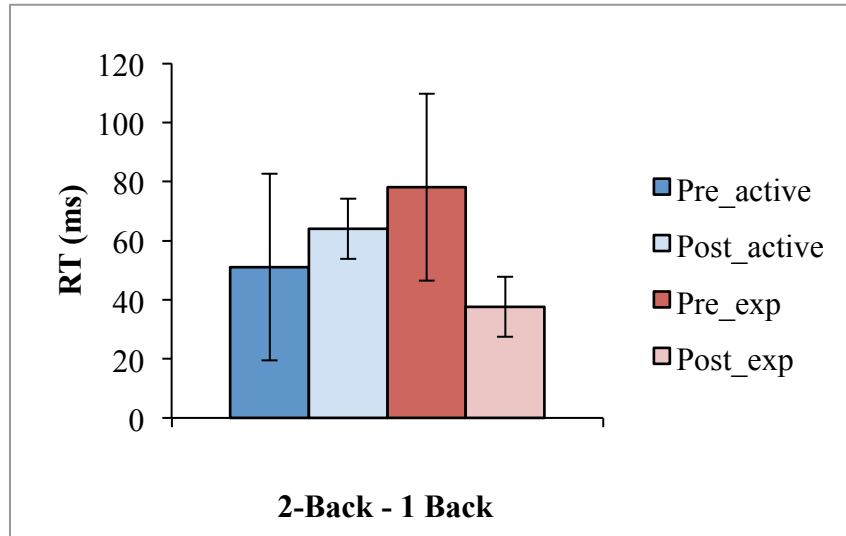


Figure 7- Figure shows the focus switching direct retrieval component at pre and post training assessment for the active control (active) and the experimental (exp) group. There were no significant main effects and interactions present. Error bars denote the standard error of the mean.

For the focus switching requiring search component (Figure 8) there was a main effect of N-Back level, $F(3, 168) = 163.13$, $p < 0.05$, $\eta^2p = 0.74$. Post hoc analysis showed a linear increase in RT as the N-Back level increased-- 2-Back ($M = 56.15$, $SD = 47.31$) vs. 3-Back ($M = 126.84$, $SD = 75.04$), $t(56) = 9.45$, $p < 0.05$, 3-Back vs. 4-Back ($M = 224.33$, $SD = 109.40$), $t(56) = 9.48$, $p < 0.05$, 4-Back vs. 5-Back ($M = 371.58$, $SD = 157.87$), $t(56) = 8.03$, $p < 0.05$. There was no main effect of time, $F(1, 56) = 1.69$, $p = 0.20$, $\eta^2p = 0.03$ or group, $F(1, 56) = 0.02$, $p = 0.89$, $\eta^2p = 0.001$. Thus, the focus switching cost remained the same even after training for both the experimental and the

active control groups, suggesting that the component cannot be optimally trained. None of the interactions were significant, all F s < 1.69, largest $\eta^2_p = 0.03$.

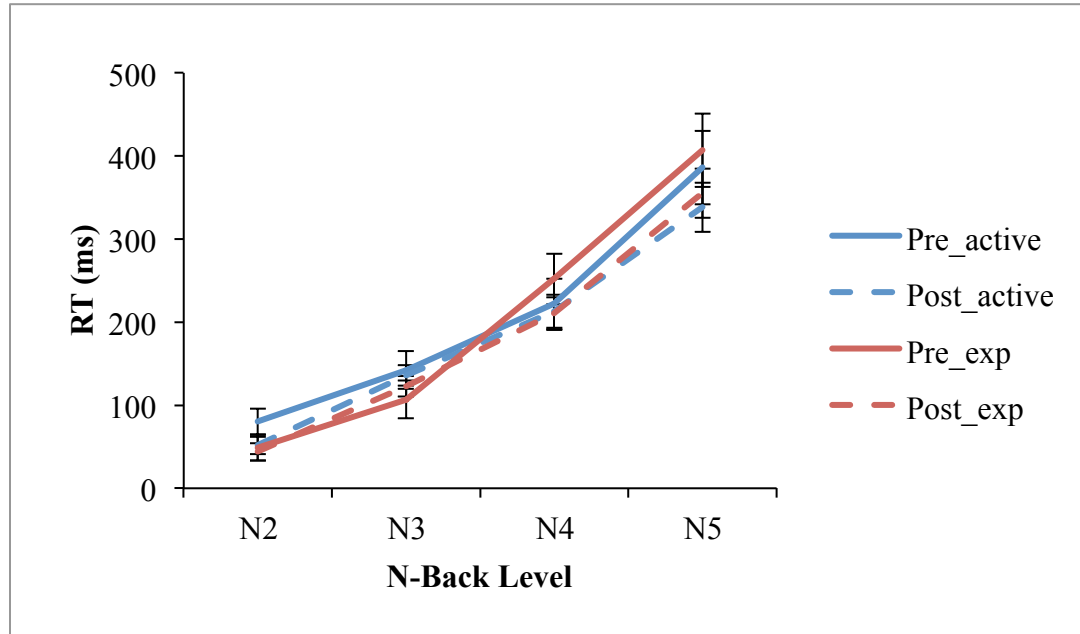


Figure 8 - Figure shows the focus switching requiring search component at pre and post training assessment for the active control (active) and the experimental (exp) group. It is the difference RT between switch and non-switch trials calculated from the random-no-updating task. Significant main effect of N-Back level was present for both the groups. Error bars denote the standard error of the mean.

For the updating component (Figure 9) we found significant main effects of time, $F(1, 56) = 13.07, p < 0.05, \eta^2_p = 0.19$, group, $F(1, 56) = 5.55, p < 0.05, \eta^2_p = 0.09$ and N-Back level, $F(4, 224) = 14.36, p < 0.05, \eta^2_p = 0.20$. Participants on an average took 493.14 ms ($SD = 322.25$) at pre assessment and 345.59 ms ($SD = 243.80$) at post assessment to update. Thus, they got faster on the updating process with training. We compared RTs of 1-Back ($M = 321.45, SD = 204.42$) with 2-Back ($M = 352.56, SD = 186.44$), 2-Back with 3-Back ($M = 416.05, SD = 208.80$), 3-Back with 4-Back ($M =$

480.38, $SD = 341.57$) and 4-Back with 5-Back ($M = 526.40$, $SD = 378.99$). After Bonferroni correction, we found significant difference only between 2-Back and 3-Back, $t(57) = 3.52$, $p < 0.05$. There was a linear increase in RT with N, however this increase was significantly larger in going from N=2 to N=3. Note that the active control group ($M = 487.03$, $SD = 225.79$) was not trained on the updating task, thus their RT was significantly higher than the experimental group ($M = 351.71$, $SD = 211.31$), $t(56) = 2.36$, $p < 0.05$. In addition, there was a significant interaction between time and group, $F(1, 56) = 13.57$, $p < 0.05$, $\eta^2p = 0.56$. The two groups had similar RT at pre training (active control: $M = 482.29$, $SD = 308.86$; experimental: $M = 503.99$, $SD = 340.24$), $t(56) = 0.25$, $p = 0.80$, however post training the experimental group ($M = 199.43$, $SD = 134.05$) had significantly lower RT compared to the active control group ($M = 491.76$, $SD = 242.44$), $t(56) = 5.68$, $p < 0.05$. In addition, there was no difference between pre ($M = 482.29$, $SD = 308.86$) and post scores ($M = 491.76$, $SD = 242.44$) for the active control group, $t(56) = 0.16$, $p = 0.88$, whereas the difference was statistically significant for the experimental group (pre: $M = 503.99$, $SD = 340.24$; post: $M = 199.43$, $SD = 134.06$), $t(56) = 5.50$, $p < 0.05$. These results show that the main effect of time is driven by whether the participants were trained on the task or not and cannot be attributed to test-retest effect. There was also a significant interaction between time and N-Back level, $F(4, 224) = 5.84$, $p < 0.05$, $\eta^2p = 0.10$. Bonferroni-corrected post hoc analysis showed significant difference between 3-Back pre ($M = 499.71$, $SD = 301.02$) and 3-Back post ($M = 332.38$, $SD = 231.04$), $t(56) = 3.78$, $p < 0.05$; 4-Back pre ($M = 574.83$, $SD = 467.77$) and 4-Back post ($M = 385.94$, $SD = 350.52$), $t(56) = 3.09$, $p < 0.05$; 5-Back pre ($M = 639.81$, $SD = 567.21$) and 5-Back post ($M = 413.00$, $SD = 358.63$), $t(56) = 3.03$, $p < 0.05$; 2-Back pre

($M = 411.23$, $SD = 293.91$) and 3-Back pre ($M = 499.71$, $SD = 301.02$), $t(56) = 2.98$, $p < 0.05$. Participants were significantly faster at post training assessment for 3-Back, 4-Back and 5-Back conditions when compared to the RTs of the same N-Back levels at pre training assessment. In addition, even though there was a linear increase in RT as the N-Back level increased at pre training assessment, this increase was significantly larger in going from 2-Back to 3-Back. Group x N-Back level interaction was also significant, $F(4, 224) = 2.58$, $p < 0.05$, $\eta^2p = 0.04$. However, after Bonferroni correction none of the post hoc comparisons were significant. The three way interaction between N-Back level, group and time was also not statistically significant, $F(4, 224) = 1.23$, $p = 0.30$, $\eta^2p = 0.02$.

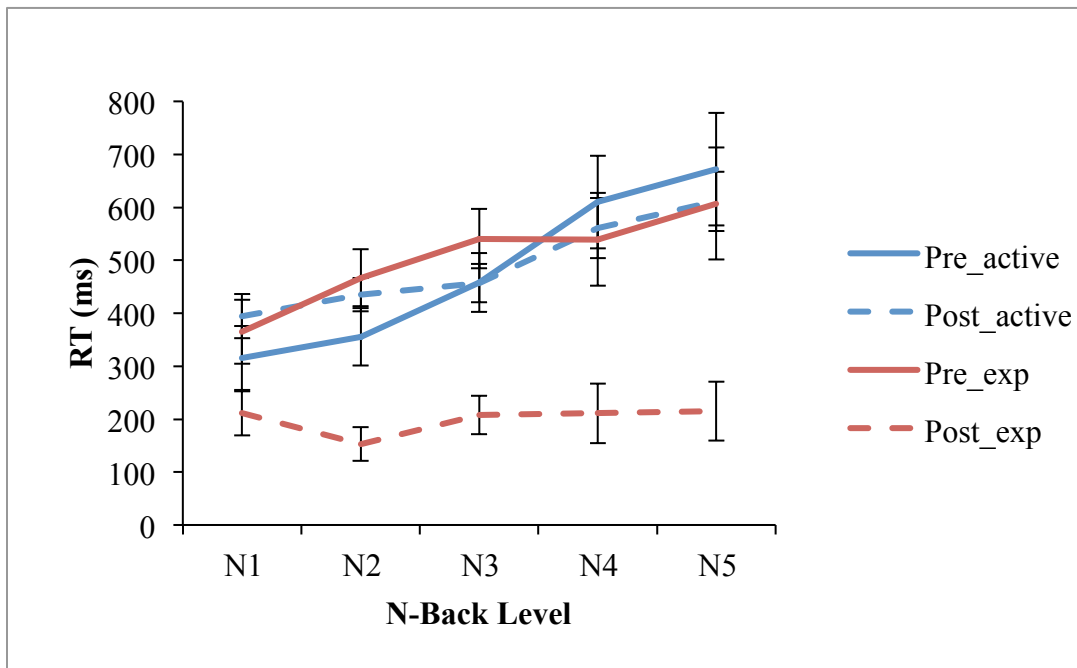


Figure 9 - Figure shows the updating component at pre and post training assessment for the active control (active) and the experimental (exp) group. It is the difference RT between forward-updating and the forward-no-updating tasks. There was a significant main effect of time, group and N-Back level. In addition time x group and time x N-Back level interactions were also significant. Error bars denote the standard error of the mean.

For the size of focus of attention (Figure 4) there was a significant main effect of time, $F(1, 56) = 281.36, p < 0.05, \eta^2p = 0.83$ and N-Back level, $F(4, 224) = 12.39, p < 0.05, \eta^2p = 0.18$. There was no main effect of group, $F(1, 56) = 0.07, p = 0.80, \eta^2p = 0.01$. Since there was a main effect of time, paired t-tests were done comparing each level of N with the next higher level of N. This was done separately for the pre and post training data. Pre training data-- after applying Bonferroni correction we found significant difference between 1-Back ($M = 677.87, SD = 200.36$) and 2-Back condition ($M = 742.44, SD = 142.75$), $t(57) = 2.90, p < 0.05$. None of the other comparisons were significant. For the post training data after Bonferroni correction we found significant difference between 1-Back ($M = 442.43, SD = 87.92$) and 2-Back condition ($M = 493.26, SD = 80.94$), $t(57) = 6.93, p < 0.05$, as well as between 2-Back and 3-Back condition ($M = 474.54, SD = 72.87$), $t(57) = 3.25, p < 0.05$. None of the other comparisons were significant. The pre training data thus yielded a step function between 1-Back and higher levels of N. The post training data too yielded a step function with a difference between 1-Back and 2-Back and again between 2-Back and all the higher levels of N. This result is not consistent with then on existing previous study, which show that after training there is a gradual increase in RT as a function of increasing N rather than a step, signifying an increase in the size of the focus of attention. It should be noted here that the RT for 2-Back at post training is higher than RT for 3-Back. It is also higher than 4-Back ($M = 472.89, SD = 74.14$) and 5-Back ($M = 475.32, SD = 65.86$) conditions. This data point might be interpreted as an anomaly since there is no theoretical explanation for a higher RT at 2-Back level compared to other higher levels of N. In this case, there was no effect

of training on the size of the focus of attention, with step at $N = 1$ at both pre and post time points.

3.3 Pre and Post Training Transfer Tasks Assessment

Eleven different tasks from five different constructs were used to evaluate the transfer effect of N-Back training. The pre training correlations between these transfer measures are shown in Table 3. Principal component analysis was run on the pre training scores of these tasks to identify and compute composite scores for the factors underlying these measures. Tasks using accuracy as the dependent measure were reversed scored. Since the constructs are theoretically correlated, oblimin rotation was used. Based on Eigen values four factors were extracted, which explained a total of 58.04% variance in the data. The factor loadings are shown in Table 4. The loadings did not conform to the theoretical grouping between the tasks and the constructs. Either the tasks representing the same constructs did not load on to the same factor (e.g. Stroop, antisaccade and plus minus) or the loading did not make theoretical sense (e.g. keep track loads negatively on to factor 1, whereas letter memory loads positively on it). Confirmatory factor analysis was also attempted, but the model did not converge. Thus, all further analysis was done at the task level rather than the proposed construct level. 2 (group: experimental and active control) x 2 (time: pre and post) repeated-measures ANOVA were conducted to see if the scores on each task changed after training. Table 5 shows the means and standard deviations for the two training groups at pre and post training for all the transfer measures.

Table 3 - Correlation between the different transfer measures at pre training.

	1	2	3	4	5	6	7	8	9	10	11
1. RAPM	--										
2. BOMAT	.42**	--									
3. Keep Track	.15	.39**	--								
4. Letter Memory	-.09	-.04	-.02	--							
5. Plus Minus	.02	.14	.09	.04	--						
6. Number Letter	.03	.29*	.07	.22	.02	--					
7. Stroop	.25	.26*	.14	-.22	.12	.08	--				
8. Antisaccade	.01	.01	-.14	.06	-.34**	.10	-.23	--			
9. Stop Signal	.16	-.11	.07	-.08	.06	-.07	-.07	.14	--		
10. Garavan	.15	.29*	.27*	.16	.21	.31*	.23	-.01	-.06	--	
11. Spatial Focus Switch	-.15	-.32*	-.10	-.20	.06	-.15	-.03	-.04	.05	-.22	--

Note: ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

Table 4 - Factor loadings based on principal component analysis with oblimin rotation for the transfer tasks at pre training.

	Factors			
	1	2	3	4
RAPM	.70			
BOMAT	.64	.43		
Number letter		.66		
Plus minus			-.82	

Table 4 (continued)

Stroop	.64		-.31
Antisaccade		.72	
Stop signal			.90
Garavan		.60	
Spatial focus switch		-.57	
Keep track	.38		
Letter memory	-.55	.66	

Note: Coefficients lower than 0.3 have not been shown here.

Table 5 - Means and standard deviations for the different transfer measures at pre and post training for the experimental and the active control group.

	Experimental		Active Control	
	Pre (<i>M/SD</i>)	Post (<i>M/SD</i>)	Pre (<i>M/SD</i>)	Post (<i>M/SD</i>)
BOMAT	60.00 (15.84)	60.46 (18.08)	70.57 (16.48)	65.52 (15.74)
RAPM	64.94 (2.32)	63.41 (16.03)	71.26 (16.93)	73.56 (12.74)
Keep Track	74.22 (12.50)	81.12 (12.17)	76.19 (11.73)	84.24 (12.01)
Letter Memory	74.18 (17.80)	85.99 (8.31)	76.22 (14.21)	78.95 (18.16)
Plus Minus*	3.6 (184.96)	98.72 (544.39)	60.52 (157.21)	105.93 (562.41)
Number Letter*	599.53 (174.98)	523.66 (173.53)	524.78 (142.99)	444.16 (130.74)
Stroop*	51.57 (61.01)	18.19 (41.81)	33.69 (47.81)	22.57 (39.91)
Antisaccade	90.77 (9.37)	94.09 (5.39)	90.99 (12.86)	94.37 (6.35)
Stop Signal	86.85 (23.52)	88.15 (28.22)	83.19 (32.22)	87.07 (30.39)
Garavan*	611.34 (248.04)	501.50 (274.29)	485.95 (261.59)	426.38 (197.70)
Spatial Focus Switch*	243.66 (218.03)	236.87 (239.12)	144.58 (203.92)	168.55 (208.73)

Note: * denotes that RT was the dependent measure.

3.3.1 Fluid Intelligence

BOMAT and RAPM (Figure 10) were used to assess transfer on this construct. For BOMAT there was a significant main effect of group, $F(1, 56) = 4.35, p < 0.05, \eta^2p = 0.07$. The active control group ($M = 68.05, SD = 13.02$) scored higher than the experimental group ($M = 60.23, SD = 15.43$). However, there was no significant effect of time, $F(1, 56) = 1.09, p = 0.30, \eta^2p = 0.02$. The interaction between time and group, $F(1, 56) = 1.57, p = 0.22, \eta^2p = 0.03$ was also not significant. Similar results were seen for RAPM. There was a significant main effect of group, $F(1, 56) = 5.44, p < 0.05, \eta^2p = 0.09$. The active control group ($M = 72.42, SD = 12.51$) scored higher than the experimental group ($M = 64.18, SD = 14.33$). Neither the main effect of time, $F(1, 56) = 0.04, p = 0.85, \eta^2p = 0.01$, nor the interaction between time and group, $F(1, 56) = 0.95, p = 0.33, \eta^2p = 0.02$ was significant. Though in both BOMAT and RAPM the active control group scored higher than the control group, there was no main effect of time. Hence, there was no transfer of training to measures of fluid intelligence.

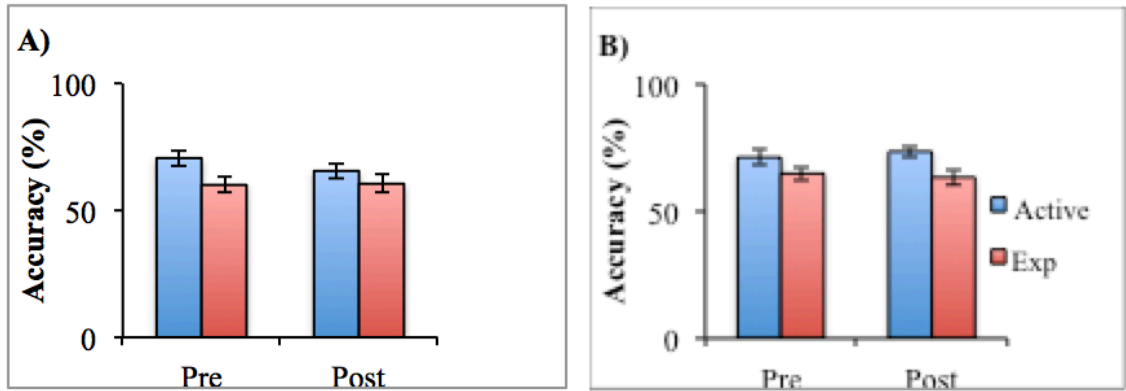


Figure 10 - Figure shows the percent accuracy score on the fluid intelligence measures before (pre) and after (post) training for the active control (active) and the experimental group (exp). Panel A- BOMAT; Panel B- RAPM. Significant effect of group was present for both the tasks. Error bars denote the standard error of the mean.

3.3.2 Updating

See Figure 11. For the keep track task there was a significant main effect of time, $F(1, 56) = 14.14, p < 0.05, \eta^2p = 0.20$, i.e. compared to pre training ($M = 75.21, SD = 12.06$), both groups got better (higher accuracy) on the task at post training assessment ($M = 82.68, SD = 12.09$). There was no main effect of group, $F(1, 56) = 1.05, p = 0.31, \eta^2p = 0.02$, and the interaction between group and time, $F(1, 56) = 1.57, p = 0.77, \eta^2p = 0.001$ was also not significant. However, for the letter memory task, there was a significant main effect of time, $F(1, 56) = 11.42, p < 0.05, \eta^2p = 0.17$ and a significant interaction between time and group, $F(1, 56) = 4.47, p < 0.05, \eta^2p = 0.07$. The main effect of group, $F(1, 56) = 0.55, p = 0.46, \eta^2p = 0.01$ was not significant. The accuracy on the task was significantly higher after training (pre: $M = 75.18, SD = 16.00$; post: $M = 82.47, SD = 14.44$). In addition, compared to the active control group ($M = 78.95, SD =$

18.16), the experimental group showed higher accuracy at post assessment ($M = 85.99$, $SD = 8.31$), $t(56) = 1.9$, $p < 0.05$. The two groups had similar accuracy before the N-Back training (active control: $M = 76.22$, $SD = 14.21$; experimental: $M = 74.14$, $SD = 17.80$), $t(56) = 0.42$, $p = 0.62$. There was no difference in accuracy before ($M = 76.22$, $SD = 14.21$) and after the training in the active control group ($M = 78.95$, $SD = 18.16$), $t(28) = 0.87$, $p = 0.39$. In contrast, the accuracy was significantly higher after training in the experimental group (pre: $M = 74.14$, $SD = 17.80$; post: $M = 85.99$, $SD = 8.31$), $t(28) = 4.02$, $p < 0.05$. Thus, there was a differential effect of training on the updating component. It lead to a very specific transfer effect (near-transfer) in the experimental group—the trained group got better on the task which was very similar (structurally and construct wise) to the task they were trained on.

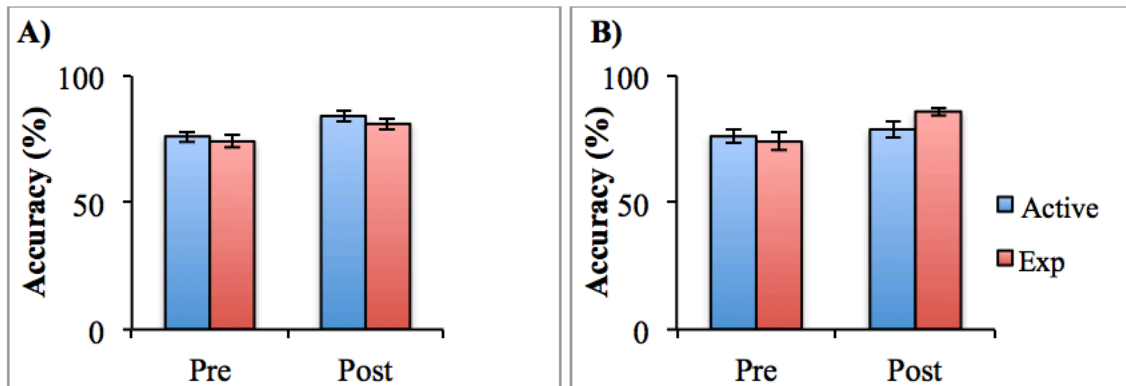


Figure 11 - Figure shows the percent accuracy score on the updating measures before (pre) and after (post) training for the active control (active) and the experimental group (exp). Panel A- Keep track; Panel B- Letter memory. Significant effect of time and time by group interaction was present for the letter memory task, whereas, only the main effect of time was significant for the keep track task. Error bars denote the standard error of the mean.

3.3.3 Task Switching

For the plus minus task there was neither a main effect of time, $F(1, 56) = 1.94, p = 0.17, \eta^2p = 0.03$, nor group, $F(1, 56) = 1.41, p = 0.24, \eta^2p = 0.03$, nor a significant interaction between time and group, $F(1, 56) = 0.76, p = 0.39, \eta^2p = 0.01$. For the number letter task there was a significant main effect of time, $F(1, 56) = 16.70, p < 0.05, \eta^2p = 0.23$. There was also a significant main effect of group, $F(1, 56) = 4.48, p < 0.05, \eta^2p = 0.07$. However, the interaction between group and time, $F(1, 56) = 0.02, p = 0.90, \eta^2p = 0.001$ was not significant. The switch cost in the number letter task was much lower after training (pre: $M = 562.16, SD = 162.81$; post: $M = 483.91, SD = 157.47$), $t(57) = 4.12, p < 0.05$ and the over all switch cost was significantly lower in the active control group compared to the experimental group (active control: $M = 484.47, SD = 117.56$; experimental: $M = 561.59, SD = 157.09$), $t(56) = 2.12, p < 0.05$. Thus, both training groups improved their performance on one measure (number letter) of task switching after training (Figure 12).

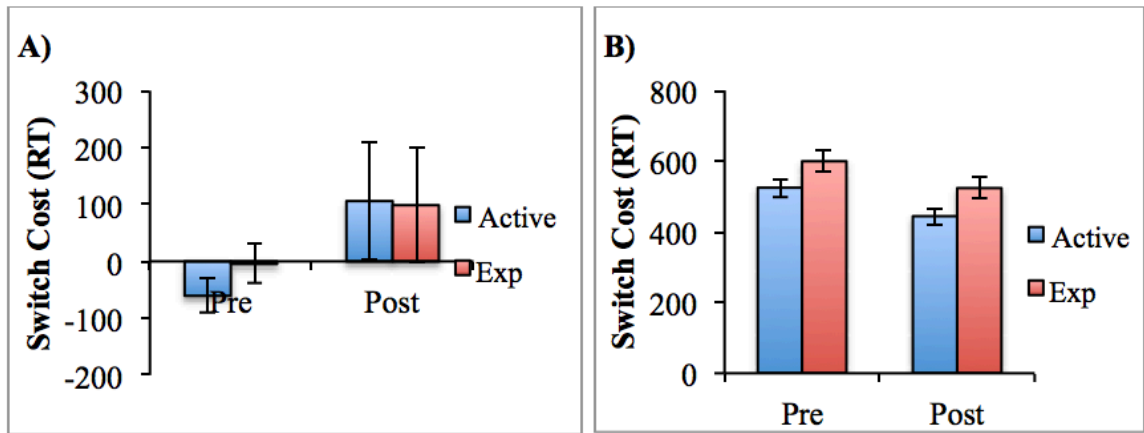


Figure 12 - Figure shows the switching cost on the task switching measures before (pre) and after (post) training for the active control (active) and the experimental group (exp). Panel A- Plus minus; Panel B- Number letter. For the plus minus task, no main effects or interaction effects were significant. For the number letter task, main effects of time and group were present. Error bars denote the standard error of the mean.

3.3.4 Inhibition

Stroop, antisaccade and stop signal tasks (Figure 13) were given to the participants to examine the effect of training on inhibition. There was a significant main effect of time, $F(1, 56) = 11.39, p < 0.05, \eta^2p = 0.17$ for the Stroop task, i.e. compared to pre training ($M = 42.63, SD = 55.07$), both groups got better (lower switch cost) on the task at post training assessment ($M = 20.38, SD = 40.57$). There was no main effect of group, $F(1, 56) = 1.35, p = 0.25, \eta^2p = 0.02$ and the interaction between group and time, $F(1, 56) = 1.32, p = 0.26, \eta^2p = 0.02$ was also not significant. Similarly, for the antisaccade task, there was a main effect of time, $F(1, 56) = 10.74, p < 0.05, \eta^2p = 0.16$ but no main effect of group, $F(1, 56) = 0.01, p = 0.88, \eta^2p = 0.001$, or a significant group x time interaction, $F(1, 56) = 0.001, p = 0.97, \eta^2p = 0.001$. The accuracy on the task significantly decreased after training (pre: $M = 90.88, SD = 11.15$; post: $M = 94.23, SD = 5.84$). Stop signal task showed neither the main effects of time, $F(1, 56) = 0.57, p = 0.46, \eta^2p = 0.01$, nor group, $F(1, 56) = 0.12, p = 0.73, \eta^2p = 0.002$, nor a significant interaction between them, $F(1, 56) = 0.14, p = 0.71, \eta^2p = 0.003$. Thus, two tasks of inhibition showed improvement with training for both training groups.

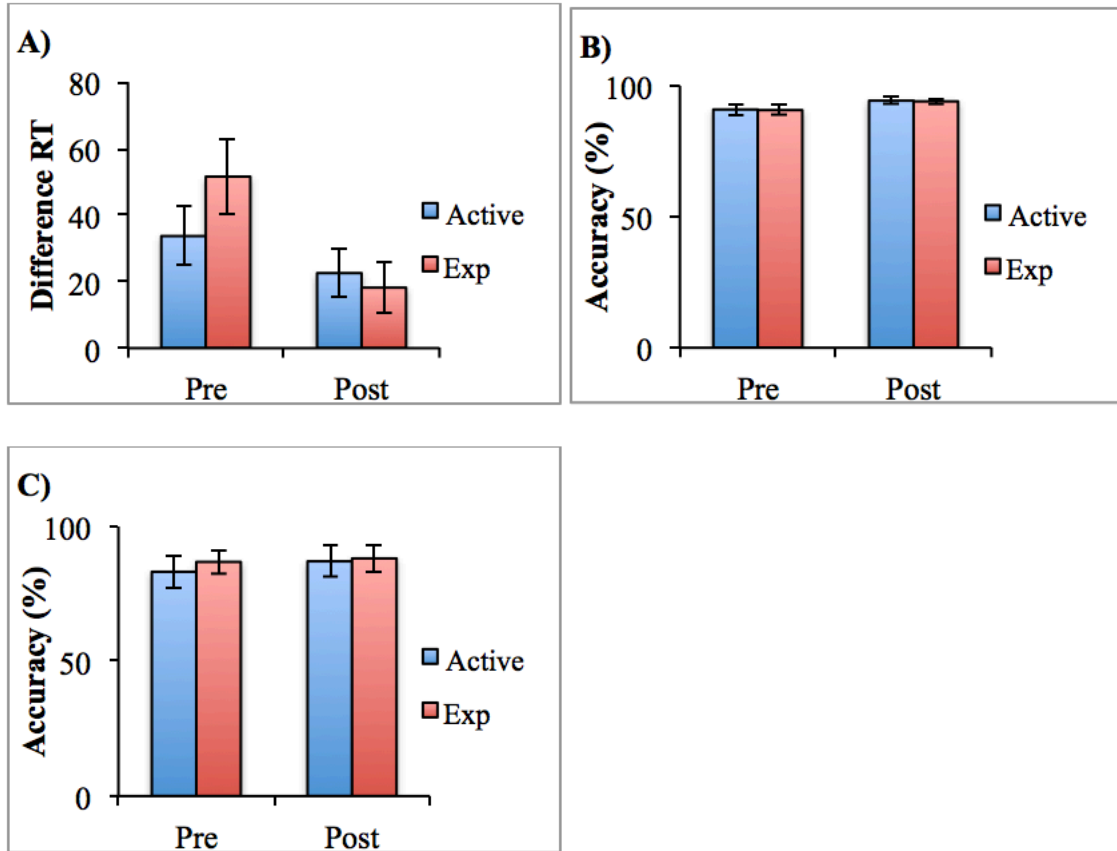


Figure 13 - Figure shows the before (pre) and after (post) training scores for the active control (active) and the experimental group on the measures of inhibition. Panel A shows the difference RT between incongruent and neutral trials in the Stroop task; Panel B shows the percent accuracy for the antisaccade task; Panel C shows the percent accuracy for the stop signal task. Significant effect of time was present for the Stroop and antisaccade tasks, no main effects and interactions were significant for the stop signal task. Error bars denote the standard error of the mean.

3.3.5 Focus Switching

On the Garavan task, we saw a main effect of time $F(1, 56) = 11.10, p < 0.05, \eta^2p = 0.17$, but no effect of group, $F(1, 56) = 2.82, p = 0.10, \eta^2p = 0.05$ or a significant group x time interaction, $F(1, 56) = 0.98, p = 0.33, \eta^2p = 0.02$. The participants were

significantly faster on this task after training (pre: $M = 548.65$, $SD = 260.46$; post: $M = 463.94$, $SD = 239.99$).). Spatial focus switching task showed neither the main effects of time, $F(1, 56) = 0.08$, $p = 0.78$, $\eta^2p = 0.001$, nor group, $F(1, 56) = 2.99$, $p = 0.09$, $\eta^2p = 0.05$, nor a significant interaction between them, $F(1, 56) = 0.25$, $p = 0.62$, $\eta^2p = 0.005$. For the focus switching construct (Figure 14), we found that the training improved the scores on the Garavan task, but did not change anything on the spatial focus switching task.

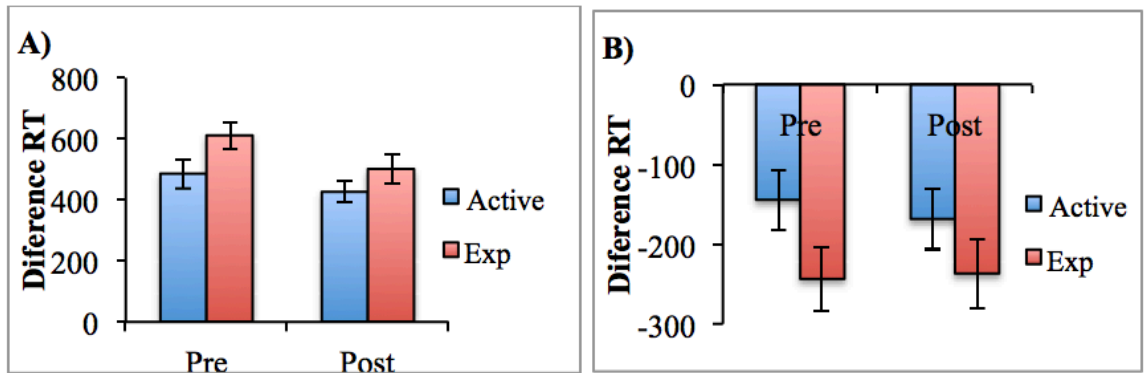


Figure 14 - Figure shows the switching cost on the focus switching measures before (pre) and after (post) training for the active control (active) and the experimental group (exp). Panel A- Garavan; Panel B- Spatial focus switching. Main effect of time was significant for the Garavan task, no significant effects were present for the spatial focus switching task. Error bars denote the standard error of the mean.

3.4 Relation Between Change in N-Back Components and Transfer Measures

To examine the main question of this proposal -- how training leads to transfer effects, correlation was run between change in the N-Back components and change in the various transfer measures, as well as the pretest score on the TOCA. Standardized change scores ((value-mean)/standard deviation) were used.

TOCA consists of eight questions, four questions are related to believing in the malleability of intelligence and the other four are related to believing that intelligence is fixed. To make the scale unidirectional, half of the questions were reverse scored such that a high total score suggested a strong belief in the malleability of intelligence.

Change in updating, focus-switching direct retrieval and TOCA did not correlate with change in any of the transfer measures. Change in focus switching requiring search correlated with change in the stop signal task (Table 6).

Table 6 - Correlation between change in the N-Back components, TOCA and change in the transfer measures.

	Updating	Focus switch requiring search	Focus switch direct retrieval	TOCA
RAPM	.04	-.09	.15	-.02
BOMAT	.01	-.14	.01	-.06
Number letter	-.13	.15	.03	-.01
Plus minus	-.15	-.08	-.15	-.12
Stroop	.02	.04	-.04	-.08
Antisaccade	-.09	-.12	-.06	-.10
Stop signal	.07	.31*	.09	.04
Garavan	-.06	-.08	-.19	.04
Spatial focus switching	.00	-.02	.12	-.19
Keep track	-.03	-.12	-.12	-.07
Letter memory	.00	.23	-.05	.03

Note: * Correlation is significant at the 0.05 level (2-tailed).

3.5 Strategy Questionnaire

Use of specific strategies to complete the training tasks was also evaluated. The results from the open-ended questionnaire suggested nine types of strategies that participants used: chunking, memorizing, visualizing, pattern identification, episodic association, recollection, anchoring, identifying a rhythm and speed. Chunking was defined as combining numbers (participants usually combined 2-5 digits) to form one unit of information; memorizing was defined as using rote learning; visualizing was defined as making a pattern either on the keyboard or using mental imagery; pattern identification was when participants found a pattern in the numbers to memorize, such as a series of odd numbers; episodic association was when participants associated a series of numbers to personally relevant information, like the birth year of their father; recollection was when participants recalled the whole series with each probed number; anchoring was when participants focused on numbers at certain locations which acted like anchors to recall the other information, e.g. memorizing the first/last digit or the middle digit in the series; identifying a rhythm, was when participants associated rhythm/beats with the numbers; speed was when participants did the task as fast as possible. Here are sample responses which participants gave which were then coded under the nine categories. Chunking- “I remembered the numbers in groups of 3-4, and remembered the order these groups were in”; Memorizing- “I memorized a sequence of 8 numbers and tried to group the rest into familiar sequences, like a descending group (9876)”; Visualizing- “In this task, I just repeated the numbers mentally 7-9 times. First, I grouped them into 3 or 4 to make memorizing easier. After a while, I began to use the visualization of pressing the number out on the keypad as a way of memorizing”; Pattern identification- “if there is a pattern where the numbers increase/decrease in order I group these numbers all together

depending on the numbers. For example: 9876 (group of 4) or 23654 (group of 5). Even if the numbers aren't exactly in order but the group contains all numbers from 2-6 as in the second example, I group them together"; Episodic association- "Sometimes I would recognize a number from somewhere in the past (an old address, for example) and remember the corresponding bundle of digits by association"; Recollection- "Yes, remember the first numbers and run through them each time a new number comes up and if it's a new number replace it with the old number and run through the current list again"; Anchoring- "I also made sure to memorize the first and last numbers in the set well because this helped me remember the numbers in the middle. In sets where I could not easily identify a pattern, it was harder to memorize"; Rhythm identification- "Yes, as I said the numbers to myself in my mind I tried to find a rhythm to the digits, sometimes it was pairs of two digits that I memorized at a time, sometimes three"; Speed- "Yes, Did it fast so I don't confuse new numbers with the older once".

Data were combined across sessions 1 and 3, 5 and 7, and 9 and 10, and dummy-coded. A 9 (strategies) x 3 (time: session 1/3, session 5/7, session 9/10) x 2 (group: experimental and active control) x 2 (task: forward-no-updating and random-no-updating) repeated-measures ANOVA was performed to see if there was any difference in use of strategy based on task, time or group. To compare differences in use of strategies for each task within group a 9 (strategies) x 3 (time: session 1/3, session 5/7, session 9/10) x 3 (task: forward-no-updating, random-no-updating and forward-updating) and a 9 (strategies) x 3 (time: session 1/3, session 5/7, session 9/10) x 2 (task: forward-no-updating and random-no-updating) repeated-measures ANOVA was conducted for the experimental and active control group respectively.

Figure 15 represents strategy use over time for each group. A 9 (strategies) x 3 (time: session 1/3, session 5/7, session 9/10) x 2 (group: experimental and active control) x 2 (task: forward-no-updating and random-no-updating) repeated measures ANOVA showed main effects of strategies, $F(8, 448) = 59.21, p < 0.05, \eta^2_p = 0.51$. Follow up t -tests were performed and p value was Bonferroni corrected. Chunking ($M = 0.79, SD = 0.25$) was significantly more often used compared to all the other strategies; memorization ($M = 0.49, SD = 0.27$), $t(57) = 5.21, p < 0.05$, visualizing ($M = 0.21, SD = 0.26$), $t(57) = 13.0, p < 0.05$, pattern identification ($M = 0.43, SD = 0.34$), $t(57) = 6.61, p < 0.05$, episodic association ($M = 0.29, SD = 0.33$), $t(57) = 10.7, p < 0.05$, recollection ($M = 0.16, SD = 0.23$), $t(57) = 13.31, p < 0.05$, speed ($M = 0.03, SD = 0.1$), $t(57) = 21.15, p < 0.05$, anchoring ($M = 0.07, SD = 0.18$), $t(57) = 17.47, p < 0.05$ and identifying rhythm ($M = 0.08, SD = 0.2$), $t(57) = 16.8, p < 0.05$. Memorization was significantly more often used than visualizing, $t(57) = 5.95, p < 0.05$, recollection, $t(57) = 7.73, p < 0.05$, speed, $t(57) = 13.01, p < 0.05$, anchoring, $t(57) = 10.49, p < 0.05$ and rhythm identification, $t(57) = 9.75, p < 0.05$. Visualizing was more often used compared to speed only, $t(57) = 4.42, p < 0.05$, whereas, pattern identification was more often used than visualizing, $t(57) = 4.06, p < 0.05$, recollection, $t(57) = 5.53, p < 0.05$, speed, $t(57) = 8.26, p < 0.05$, anchoring, $t(57) = 7.86, p < 0.05$, and rhythm identification, $t(57) = 6.76, p < 0.05$. Episodic association was more often used than speed, $t(57) = 5.61, p < 0.05$, anchoring, $t(57) = 4.3, p < 0.05$ and rhythm identification, $t(57) = 4.3, p < 0.05$, whereas, recollection strategy was used more often than speed, $t(57) = 3.92, p < 0.05$ and anchoring, $t(57) = 2.66, p < 0.05$.

There was significant interaction between strategies and time, $F(16, 896) = 5.4, p < 0.05, \eta^2p = 0.09$. Over the sessions, participants started favoring some strategies more than the other. After applying Bonferroni correction it was seen that chunking was more often used during sessions 9/10 ($M = 0.89, SD = 0.27$) compared to sessions 1/3 ($M = 0.71, SD = 0.38$), $t(57) = 3.4, p < 0.05$ and sessions 5/7 ($M = 0.78, SD = 0.38$), $t(57) = 2.35, p < 0.05$. Memorization was used less often in sessions 5/7 ($M = 0.41, SD = 0.42$), $t(57) = 3.4, p < 0.05$ and sessions 9/10 ($M = 0.4, SD = 0.42$), $t(57) = 3.45, p < 0.05$ compared to sessions 1/3 ($M = 0.66, SD = 0.38$). Episodic associations were made more often in sessions 9/10 ($M = 0.36, SD = 0.43$) compared to sessions 1/3 ($M = 0.18, SD = 0.35$), $t(57) = 3.24, p < 0.05$, whereas speed was also more often used in session 9/10 ($M = 0.36, SD = 0.43$) compared to sessions 1/3 ($M = 0.03, SD = 0.11$), $t(57) = 5.78, p < 0.05$ and sessions 5/7 ($M = 0.03, SD = 0.11$), $t(57) = 5.78, p < 0.05$. The interaction between strategies and tasks was also statistically significant, $F(8, 448) = 2.84, p < 0.05, \eta^2p = 0.05$. However, none of the follow-up t -tests were significant after Bonferroni correction. There was also a significant three way interaction between strategies, task and group, $F(8, 448) = 1.98, p < 0.05, \eta^2p = 0.03$. Again, none of the follow-up t -tests were significant after Bonferroni correction. In addition, there was a main effect of group, $F(1, 56) = 5.24, p < 0.05, \eta^2p = 0.09$. The active control group used more strategies in the forward-no-updating and random-no-updating tasks ($M = 0.31, SD = 0.10$) compared to the experimental group ($M = 0.25, SD = 0.08$).

To compare the use of strategies for each task within each training group a 9 (strategies) x 3 (time: session 1/3, session 5/7, session 9/10) x 3 (task: forward-no-updating, random-no-updating and forward-updating) repeated-measures ANOVA was

conducted for the experimental group, and a 9 (strategies) x 3 (time: session 1/3, session 5/7, session 9/10) x 2 (task: forward-no-updating and random-no-updating) repeated-measures ANOVA was performed for the active control group. In the experimental group, a significant main effect of tasks emerged, $F(2, 56) = 7.1, p < 0.05, \eta^2p = 0.2$ and strategies, $F(8, 224) = 26.05, p < 0.05, \eta^2p = 0.48$. After applying Bonferroni correction it was seen that more strategies were employed to complete the forward-no-updating task ($M = 0.27, SD = 0.10$) than the forward-updating task ($M = 0.20, SD = 0.08$), $t(28) = 3.57, p < 0.05$. There was no significant difference in the number of strategies used while doing the forward-no-updating and random-no-updating tasks ($M = 0.24, SD = 0.08$), $t(28) = 1.32, p = 0.20$ or the forward-updating and the random-no-updating tasks, $t(28) = 2.51, p = 0.02$. Chunking ($M = 0.43, SD = 0.30$) was significantly more often used than visualizing ($M = 0.07, SD = 0.14$), $t(28) = 6.24, p < 0.05$, pattern identification ($M = 0.19, SD = 0.23$), $t(28) = 3.64, p < 0.05$, episodic association ($M = 0.13, SD = 0.22$), $t(28) = 6.19, p < 0.05$, anchoring ($M = 0.06, SD = 0.16$), $t(28) = 6.12, p < 0.05$ and rhythm identification ($M = 0.09, SD = 0.21$), $t(28) = 5.84, p < 0.05$. Memorization ($M = 0.62, SD = 0.30$) technique was more often used compared to visualizing, $t(28) = 8.07, p < 0.05$, pattern identification, $t(28) = 5.77, p < 0.05$, episodic association, $t(28) = 6.03, p < 0.05$, recollection ($M = 0.19, SD = 0.26$), $t(28) = 7.38, p < 0.05$, speed ($M = 0.16, SD = 0.27$), $t(28) = 6.61, p < 0.05$, anchoring, $t(28) = 9.39, p < 0.05$ and rhythm identification, $t(28) = 6.71, p < 0.05$. There was also a significantly higher use of the recollection strategy compared to rhythm identification, $t(28) = 1.6, p < 0.05$.

There was a significant interaction between tasks and strategies used, $F(16, 448) = 8.9, p < 0.05, \eta^2p = 0.24$. Post hoc comparisons after Bonferroni correction showed that

there was a difference in use of strategies based on task between forward-no-updating and forward-updating as well as between random-no-updating and forward-updating. In comparison to the forward-updating task ($M = 0.33$, $SD = 0.37$), participants used chunking more often while doing the forward-no-updating task ($M = 0.77$, $SD = 0.28$), $t(28) = 5.38$, $p < 0.05$, and the random-no-updating task ($M = 0.70$, $SD = 0.34$), $t(28) = 4.51$, $p < 0.05$. Similarly, they used pattern identification more often while completing the forward-no-updating task ($M = 0.47$, $SD = 0.40$), $t(28) = 4.42$, $p < 0.05$ and the random-no-updating task ($M = 0.45$, $SD = 0.41$), $t(28) = 3.45$, $p < 0.05$, compared to the forward-updating task ($M = 0.10$, $SD = 0.28$). Memorizing strategy was relied more heavily upon while completing the forward-updating task ($M = 0.70$, $SD = 0.35$) compared to the forward-no-updating ($M = 0.43$, $SD = 0.34$), $t(28) = 4.04$, $p < 0.05$ and the random-no-updating task ($M = 0.36$, $SD = 0.31$), $t(28) = 5.13$, $p < 0.05$. Participants also used speed more often in doing the forward-updating task task ($M = 0.21$, $SD = 0.36$) compared to the forward-no-updating ($M = 0.01$, $SD = 0.01$), $t(28) = 3.09$, $p = 0.005$ and the random-no-updating task ($M = 0.01$, $SD = 0.06$), $t(28) = 3.00$, $p = 0.006$. However, the t-tests for this comparison failed to reach statistical significance after Bonferroni correction. There was also a significant interaction between tasks and time, $F(4, 112) = 2.56$, $p < 0.05$, $\eta^2p = 0.08$. None of the t -tests were significant after Bonferroni correction. Interaction between strategies and time was also significant, $F(16, 448) = 5.07$, $p < 0.05$, $\eta^2p = 0.15$. Participants increased/decreased the use of specific strategies over sessions. Compared to sessions 1/3 ($M = 0.48$, $SD = 0.33$) chunking was used more often by session 9/10 ($M = 0.71$, $SD = 0.26$), $t(28) = 3.46$, $p < 0.05$. However, use of

memorizing strategy was decreased when comparing sessions 1/3 ($M = 0.68$, $SD = 0.33$) to sessions 9/10 ($M = 0.36$, $SD = 0.33$), $t(28) = 4.41$, $p < 0.05$.

In the active control group, a significant main effect of strategies emerged, $F(8, 224) = 34.98$, $p < 0.05$, $\eta^2p = 0.56$. After applying Bonferroni correction it was seen that, except memorizing ($M = 0.59$, $SD = 0.25$) $t(28) = 3.48$, $p = 0.003$, chunking ($M = 0.84$, $SD = 0.22$) was significantly more often used than all the other strategies -- visualizing ($M = 0.26$, $SD = 0.27$), $t(28) = 9.01$, $p < 0.05$, pattern identification ($M = 0.40$, $SD = 0.31$), $t(28) = 6.38$, $p < 0.05$, episodic association ($M = 0.33$, $SD = 0.34$), $t(28) = 7.86$, $p < 0.05$, recollection ($M = 0.13$, $SD = 0.24$), $t(28) = 11.13$, $p < 0.05$, speed ($M = 0.05$, $SD = 0.14$), $t(28) = 15.63$, $p < 0.05$, anchoring ($M = 0.07$, $SD = 0.19$), $t(28) = 13.67$, $p < 0.05$ and rhythm identification ($M = 0.1$, $SD = 0.23$), $t(28) = 11.27$, $p < 0.05$. Memorization technique was more often used compared to visualizing, $t(28) = 4.87$, $p < 0.05$, recollection, $t(28) = 7.62$, $p < 0.05$, speed, $t(28) = 11.54$, $p < 0.05$, anchoring, $t(28) = 8.93$, $p < 0.05$ and rhythm identification, $t(28) = 8.81$, $p < 0.05$. There was also a significantly higher use of the pattern strategy compared to recollection, $t(28) = 4.85$, $p < 0.05$, speed, $t(28) = 5.13$, $p < 0.05$, anchoring, $t(28) = 5.05$, $p < 0.05$ and rhythm identification, $t(28) = 3.92$, $p < 0.05$. In addition, episodic association was more often used compared to speed, $t(28) = 3.87$, $p < 0.05$.

There was a significant interaction between tasks and strategies used, $F(8, 224) = 3.32$, $p < 0.05$, $\eta^2p = 0.11$. However, none of the t -tests were significant after Bonferroni correction. There was also a significant interaction between strategies and time, $F(16, 448) = 1.98$, $p < 0.05$, $\eta^2p = 0.07$. Participants increased the use of speed in sessions 9/10

($M = 0.40$, $SD = 0.43$) compared to sessions 1/3 ($M = 0.03$, $SD = 0.13$), $t(28) = 4.42$, $p < 0.05$ and sessions 5/7 ($M = 0.05$, $SD = 0.15$), $t(28) = 4.00$, $p < 0.05$.

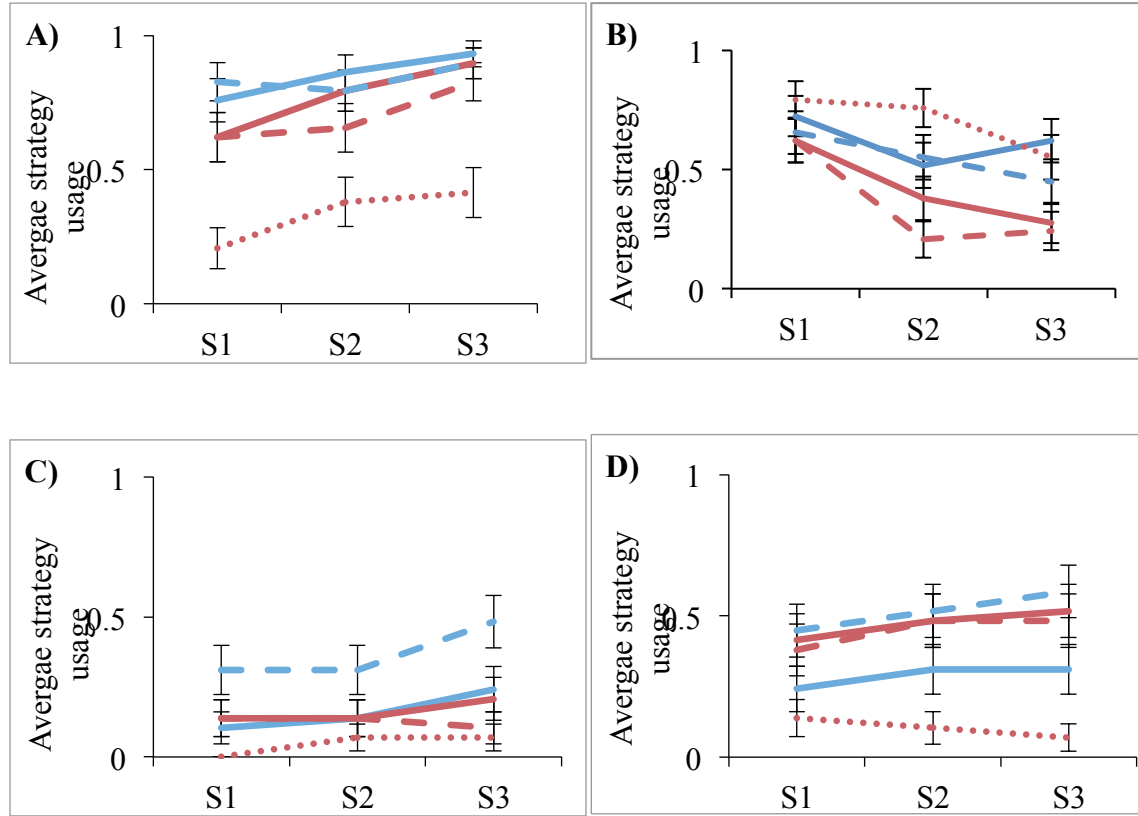


Figure 15 - The figure shows the use of nine different strategies reported by the experimental group (exp) and the active control (active) group to perform the forward-no-updating (FWNU), random-no-updating (RNU) and the forward-updating (FWU) tasks across the ten training sessions. S1- Sessions 1/3, S2- Sessions 5/7, S3- Sessions 9/10; Panel A- chunking, Panel B- memorizing, Panel C- visualizing, Panel D- pattern recognition, Panel E- episodic association, Panel F- recollection, Panel G- speed, Panel H- anchoring, Panel I- rhythm identification. Error bars denote the standard error of the mean.

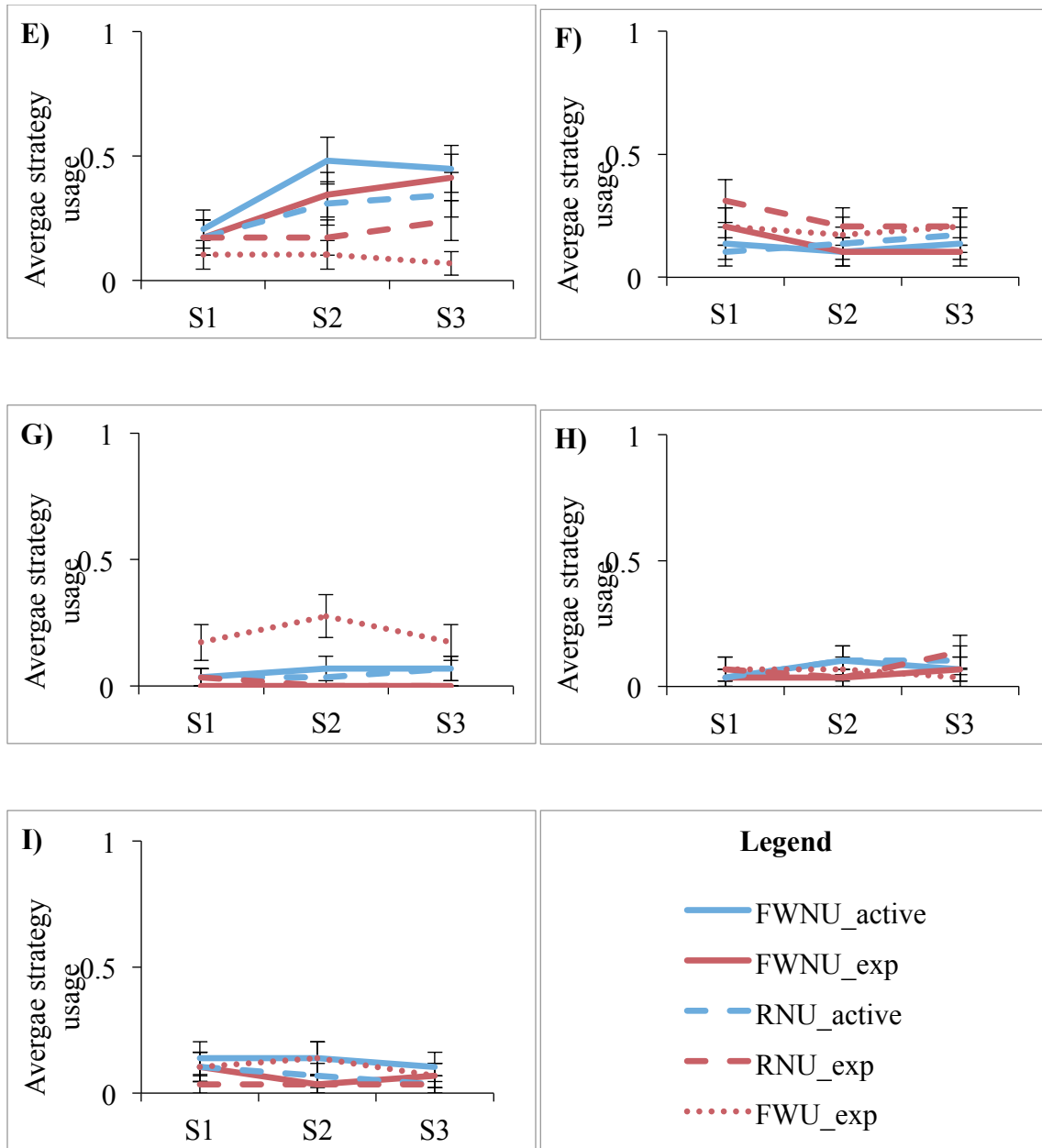


Figure 15 continued.

To see how the use of strategy influenced working memory performance, correlation was run between the absence or presence of each strategy, as well as the total number of strategies used during sessions 9/10 and the average level of N reached during the same sessions (Table 7). Performance on the random-no-updating task significantly

correlated with the pattern recognition strategy ($r = .45$) as well as with the total number of strategies used ($r = .26$). Performance on the forward-updating task was positively correlated with the chunking strategy ($r = .54$) and rhythm identification strategy ($r = .40$), and negatively correlated with the memorization strategy ($r = -.36$). Participants in the forward-updating task relied more on memorization than any other strategy and that may have actually hindered their learning progress rather than helping them. No correlations were significant for the forward no-updating condition. In addition, multiple linear regressions were run to see if using different strategies had an effect on the task performance. The regression was not significant for the forward-no-updating task, $F(9, 48) = 0.81$, $p = 0.61$, $R^2 = 0.13$. However, significant regression was found for the random-no-updating task, $F(9, 48) = 2.05$, $p < 0.05$, $R^2 = 0.28$ as well as the forward-updating task, $F(9, 19) = 2.70$, $p < 0.05$, $R^2 = 0.56$. Pattern recognition was a significant predictor for the random-no-updating task ($\beta = 0.65$, $p < .05$), whereas, chunking was a significant predictor for the forward-updating task ($\beta = 3.36$, $p < .05$). Thus, participants benefited by employing the pattern recognition strategy on the random-no-updating task and chunking strategy on the forward-updating task.

Table 7 - Correlation between the different strategies used during training sessions 9/10 and the average level of N reached during the same sessions.

	FWNU	RNU	FWU
Chunking	.03	.03	.54**
Memorizing	-.10	-.07	-.39*

Table 7 (continued)

Visualizing	.15	.24	-.06
Pattern Recognition	.20	.45**	-.32
Episodic Association	.21	.23	-.16
Recollection	.00	-.16	-.12
Speed	.07	-.11	.05
Anchor	.13	.03	-.05
Rhythm Identification	.12	.01	.40*
Total Strategies	.25 ¹	.26*	.00

Note: ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed). ¹ The correlation was close to significance, $p = 0.06$.

CHAPTER 4. DISCUSSION

This study tested a theoretically meaningful mechanism to explain transfer effects after working memory training in younger adults. A componential analysis of the modified N-Back task (Price et al., 2014; Verhaeghen et al., 2004) was conducted to understand if and how specific sub components of the training task are affected by training and how they are related to transfer effects seen on the untrained tasks. The three components included were updating, direct focus switching, and focus switching requiring search. Since updating has been found to be associated with both working memory (e.g., Miyake et al., 2000) and fluid intelligence (e.g., Friedman et al., 2006), the study specifically looked at the role of the updating component in eliciting transfer to measures of fluid intelligence. There were three straightforward main findings: (a) The updating component is more malleable than the focus switching requiring search and focus switching direct retrieval and thus may be more important in working memory interventions; (b) training in the updating aspect of the N-Back task did not generalize to other cognitive control processes implicated in working memory, namely shifting and inhibition (Miyake et al., 2000), and; (c) training the updating component did not lead to transfer to measures of fluid intelligence.

4.1 Learning Curves for the Trained Tasks

After ten hours of training, both groups showed marked improvement in their performance on tasks that they were trained on. These gains were modeled using a negative exponential curve (Gaschler, R., Progscha, J., Smallbone, K., Ram, N., & Bilalić, M., 2014). Gains can be considered under three guises: the learning slope, which

reflect the efficiency of learning, the amount of gain during training, and the asymptote, which indicates its final efficacy. Learning slopes were different across tasks. In general, the forward updating condition separated out from the other conditions, with a lower asymptote (forward-updating: experimental: 8.42; forward-no-updating: experimental: 12.27, active control: 12.65; random-no-updating: experimental: 12.17, active control: 12.68), smaller training gain (forward-updating: experimental: 2.2; forward-no-updating: experimental: 11.86, active control: 9.56; random-no-updating: experimental: 4.33, active control: 13.08), and slower learning rates (forward-updating: experimental: 0.19; forward-no-updating: experimental: 1.42, active control: 1.06; random-no-updating: experimental: 0.60, active control: 1.55; note the one exception for random no-updating in the experimental group, where the 95% CI overlapped with that of forward updating). These results suggest that the updating process itself is more difficult to perfect, on all three parameters, than the two search processes (one-step search and random order search) associated with the no-updating tasks.

In the random-no-updating condition, the learning slope was shallower for the experimental group than the active control group, even though the asymptote was similar (experimental: $c = 0.60$; active control: $c = 1.55$). There could be at least two reasons for this. First, this may be due to the anticipation or fatigue of doing three tasks for the experimental group compared to two tasks for the active control group. Second, employing more strategies may have led to faster learning for the active control group compared to the experimental group. Results showed that the active control group employed more strategies than the experimental group and higher strategy use was correlated with better performance on the random-no-updating task. We did not see a

group effect in the forward no-updating task, suggesting no effect of fatigue or anticipation in this condition. It may be the case that the random-no-updating task is cognitively more demanding than the forward-no-updating task, and that the participants in the experimental group were therefore more motivated to perform, or simply less fatigued.

Note that in the present design, the maximum level of N participants could reach was 13. This design feature thus limited asymptotic performance. Near perfect asymptotic performance was seen for the forward-no-updating task as well as the random-no-updating task. In addition, for the two non-updating tasks, this near-perfect asymptote was reached by session 2 in both groups, showing very fast learning. One implication of this ceiling effect is that this makes it impossible to investigate differences in learning rate or final performance between these two conditions. It is critical to note that the learning curves were calculated using the group average and thus, they do not capture the between-subject variability. Averaging across subjects does not provide information about the training related gains (associated with underlying cognitive change) occurring at the individual level.

4.2 Performance on the Trained Tasks

On the forward-no-updating task participants in both training groups had faster RTs for all the levels of N after training compared to before. Training effects were equally large in both groups, as would be expected, given that they were both trained on this particular task in equal amounts. Both in pretest and posttest, RT increased in going from $N = 1$ to $N = 2$. Thus, with increased task load, increased effort was required to do

the task. Surprisingly, there was a decrease in RT in going from $N = 2$ to $N = 3$, whereas there was no difference in going from $N = 3$ to $N = 4$ or from $N = 4$ to $N = 5$. In a recent meta-analysis, Bopp and Verhaeghen (2018) showed that for N larger than one there is a switch cost (higher RTs and lower accuracy) involved in bringing the items into the focus of attention. This focus switch cost increases linearly with N . Thus, the decrease in RT with increasing load (in going from $N = 2$ to $N = 3$) seen here is an anomaly, as there is no theoretical explanation for this result, and it falls outside the empirical norm for N-Back studies. The result was similar between the experimental and the active control group.

For the random-no-updating task, participants in both training groups improved their performance on both switch and non-switch trials. Training effects on RT were equally large in both groups, as would be expected, given that they were both trained on this particular task in equal amounts. In addition, RTs increased as the level of N increased, suggesting an increased task demand imposed by higher levels of N . Before the training, the participants took longer to perform the task as the level of N increased. This increase in RT was linear and seen for both switch and non-switch trials. However, after training, this linear increase in RT was only present for the switch trials, suggesting near-perfect efficiency after training for the non-switch trials. Such perfect efficiency would be expected if participants keep the last digit active within the focus of attention. The result then suggests that this is not the strategy employed before training, where a slope over N is present, probably because participants move their focus away, just as they do (and are required to do) in the forward-no-updating condition, and hence need to backtrack on a proportion of trials. The linear increase with N in switch trials, coupled

with perfect efficiency in the non-switch trials after training, suggests a Sternbergian search process through the outer store (Sternberg, 1966); the value of the slope indicates the duration of each step in the search, which is about 100 ms after training (down from 150 ms/N before training). The results suggest that this Sternbergian search process is less malleable to training than the non-switch trials. This pattern of result was seen in both training groups. Thus, training the experimental group on the updating component did not lead to transfer on the performance of the forward-no-updating task or the random-no-updating task. This, in turn, suggests independence between the updating and the switching processes.

In the forward-updating task – the one task the control group was not trained on -- the performance of the active control and the experimental group was similar before the training. There was a decrease in RT at post training compared to pre training for both groups, but, as expected, the decrease was significantly larger for the experimental group compared to the active control group. Also, the RT increased in going from 1-Back to 4-Back level (slope of about 90 ms/N before training for both groups; 60 ms/N after training in the control group, and 8 ms/N in the experimental group). Thus, as the task load increased, the participants took longer to complete the task. In addition, participants were significantly faster at post training assessment for all N-Back levels compared to the pre training assessment. An interesting finding is that after training, in the trained group, updating was performed with near-perfect efficiency (8 ms/N), that is, mostly independent of task load in the outer store, as it – theoretically – should (e.g. Price et al., 2014). Do note, however, that in the ANOVA, the N by group by time interaction failed to reach significance.

4.3 Effect of Training on the N-Back Components

The focus switching direct retrieval process and the focus switching requiring search process did not show any benefits of training. Only the updating process became more efficient (and, in fact, nearly completely efficient) with practice.

For focus switching direct retrieval both training groups had similar switching cost before and after training. This finding is different from Price et al. (2014), who found a sizable decrease in switch cost after training. In their study a 119 ms difference was found between pre and post switch cost. In the present study, the active control group showed a slight increase of 13 ms in their switch cost, whereas, the experimental group showed a slight decrease of 40 ms in their switch cost. It should be noted that at session one in Price et al.'s study, the average RT for the 1-Back forward-no-updating task was close to 490 ms and for the 2-Back level was around 610 ms. Bopp and Verhaeghen (2018) too in their meta-analysis showed that the average RT for 1-Back task is around 480 ms and for the 2-Back task is around 620 ms in younger adults. In the current study, the average RT in the forward-no-updating task for the 1 and 2-Back levels at pre training assessment were 660 ms and 739 ms for the experimental group and 695 ms and 746 ms for the active control group – much higher than the values found in the meta-analysis or in Price et al. In addition, the RTs for the 1-Back and 2-Back levels in Price's study at the end of the training were around 420 ms. This is similar to the findings of the current study: RTs for the 1-Back level for the experimental and the active control group after training were 483 ms and 446 ms respectively, whereas RTs for the 2-Back for the experimental and active control were 461 ms and 489 ms. Since participants in this study as well as in Price et al.'s study had similar RTs for 1-Back and 2-Back levels after

training, it is suggestive of a uniform and consistent effect of training on the forward-no-updating task. It is plausible that the null result of training on the focus switching direct retrieval component of working memory is driven by the unusually high RTs at pre training in the present study. It is not clear why RTs were elevated at pre training. One possibility is that our participants employed more time-consuming strategies than participants in other studies, but at present this remains a conjecture.

Similar results were seen for the focus switch requiring search. The training did not improve the efficiency of the process -- both training groups showed similar RT before and after the training. This finding is also in contrast to Price et al.'s results, where a small but significant improvement on the difference RT scores (switch - non-switch random-no-updating trials) was found. In their study, the participants became 1.5 times faster on this search process even though it did not reach perfect efficiency (a significant increase in RT over N was present even after training). Compared to Price et al.'s findings, participants in this study had similar switch costs at pre training, but did not show marked improvement on the search process as a function of training. It should be noted that this null result was obtained because the participants in the current study improved their performance on both the switch and non-switch trials of the random-no-updating task to the same extent, leading to an identical switch cost before and after the training.

In contrast to the above findings, there was a significant effect of training on the updating process in the experimental group. Participants increased the efficiency of their updating process, which resulted in an overall 148 ms decrease in their updating RT. Since the active control group was not trained on this component, their final RT was

slower (by 142 ms) than the experimental group. The difference in RT before and after training for the active control group was not significant (-8 ms). In contrast, the difference was highly significant for the experimental group (315 ms). The slope over N in updating was essentially flat after training (6 ms/N). This reiterates the point made above, that after training updating was performed with near-perfect efficiency, suggesting that it is independent of task load in the outer store, as it – theoretically – should be.

In addition, this study did not find evidence for an expansion in the size of the focus of attention as a function of training. This result is again not consistent with the previous study (Price et al., 2014), which showed that after training there is a gradual increase in RT rather than a step in the forward-no-updating task as a function of increasing N, signifying an increase in the size of the focus of attention.

4.4 Transfer Effects of the N-Back Training

The results so far support the hypothesis that selective training on the updating component of working memory indeed leads to a selective change in both the speed and efficiency with which participant perform this process. The main question of this study was whether, and if so how, changes in the updating component of working memory would lead to transfer to other cognitive control processes as well as to measures of fluid intelligence.

For the updating transfer tasks, there was mixed evidence for transfer. The experimental group improved more than the control group on the letter memory task, but not on the keep track task. In the letter memory task, letters were presented one at a time and participants kept track of the last four letters, whereas, in the keep track task,

participants were required to additionally process intervening items for meaning while maintaining the information in the outer store. Thus, the letter memory task is arguably closer to the task demands of the N-Back task than the keep track task. The study thus found a very specific transfer effect of updating training to one other updating task, with a very similar task structure and similar processing demands.

On the tasks measuring task switching, inhibition and focus switching, there was no added benefit of training the participants on the updating component. Both training groups improved their performance on the number letter task but not on the plus minus task. Similarly both training groups increased their performance on the Stroop, antisaccade and the Garavan task. However, no effects of training were seen on the stop signal and the spatial focus switching task. Thus, updating training does not transfer to measures of other cognitive control processes. In addition, since participants in both training groups got better only on some measures of the constructs studied here, it suggests that the working memory training does not train the whole construct, but leads to very specific transfer effects seen on some tasks. It is plausible that training leads to the increase in the usage or efficient usage of certain strategies, which led to these findings.

Note, of course, that the main effect of time seen in some of the tasks may be due to the test-retest effect and not actually due to the working memory training. To rule out this explanation, the results of the active control and the experimental group need to be compared to those of a passive control group. If these results were to show an interaction with group in the presence of a passive control group, stronger claims can be made about the transfer effects of the working memory training.

Finally, focusing on the central question of the study, we examined whether updating is the crucial link in eliciting transfer effects on the measures of fluid intelligence. The study did not find support for this hypothesis. There was a main effect of group for both tasks, that is, the active control group scored higher on both these measures compared to the experimental group. However, there was no effect of time and hence no effect of training in either of the trained groups, and, crucially, no time by group interaction.

4.5 N-Back Components and Transfer Measures

Contrary to expectations, the updating component did not correlate with the measures of fluid intelligence or the other measures of updating at pretest. Friedman et al.'s (2000) study used keep track and letter memory tasks as updating measures and found a significant correlation between those and fluid intelligence. It is plausible that the use of a different updating task in the present study may have resulted in the lack of a significant correlation between the updating component and fluid intelligence. Even if this were the case, the present study still used the same measures for updating as Friedman et al., and so a significant correlation between the measures of fluid intelligence, keep track task and the letter memory task would be expected. However, this was not what was found. It is not clear why no significant correlation was obtained; one possibility would be range restriction in the present sample. Participants in this study did not show a lot of improvement on the post training scores compared to the pre training scores, thus the range of the change scores will be restricted.

Focus switching requiring search was significantly correlated with a single task, namely the stop signal task. However, no theory explains this relation, thus, this correlation may be spurious. Focus switching direct retrieval also did not correlate with any of the transfer measures. In addition, the TOCA scale, which measures belief in the malleability of intelligence, did not correlate with the change scores on the trained components or the transfer measures. This finding stands in contrast to the study by Jaeggi et al. (2014), which found a significant correlation between beliefs and transfer.

4.6 Relation Between Strategies and Performance

The strategies employed by the participants were broken down into nine categories -- chunking, memorizing, visualizing, pattern recognition, episodic association, recollection, speed, anchoring and rhythm identification. Out of these, chunking was the most commonly employed strategy. Memorization, episodic association and pattern identification were also heavily relied upon. As training progressed, participants started favoring chunking, episodic association, and speed, and showed less reliance on memorization. Overall, the active control group employed more strategies than the experimental group. This may be because participants in the experimental group had more training tasks to complete, and hence were less inclined to spend as much time as the active control group in deploying strategies for each task. In the experimental group, participants used more strategies to complete the forward-no-updating task compared to the forward-updating task. It may be the case that it was easier to identify an efficient strategy for the no-updating tasks compared to the updating task, which helped them to reach the asymptote for the no-updating tasks by training session 2. In addition, chunking and pattern identification were employed more during the forward-

no-updating and the random-no-updating tasks, whereas memorization and speed were favored more for the forward-updating task. The use of specific strategies correlated with performance. Better performance on random-no-updating task was related to employing the pattern recognition strategy, whereas higher performance on forward-updating task was related to using chunking more often. It would be interesting for future studies to investigate the role of explicit strategy instructions in these conditions to elicit training gains, and if such instructions would lead to broader transfer effects.

4.7 Limitations and Future Direction

It is important to emphasize here that the lack of transfer effects in this study may be due to the use of the modified N-Back task instead of the traditional N-Back task, which is employed by the majority of the working memory training studies. The traditional N-Back task is experimenter-paced and imposes higher encoding demand on the participants. It also requires participants to keep track, a requirement that is at least partially eased by the localization schema provided by the version employed here. It is plausible that the two N-Back tasks differ in the employment of the three componential processes discussed in this study, and that this drives the presence or absence of transfer effects. More research would be necessary to concretely rule out absence of transfer effects due to the specific nature of the modified N-Back task.

In addition, the control group in this study was trained on a task that requires a high degree of cognitive control (focus switching). It is thus plausible that possible transfer effects are masked here (absence of group x time interaction), because the training on focus switching itself might have produced transfer effects. It is then essential

to compare these results with a passive control group, to rule out this possibility. Also, far transfer effects seen in other studies are usually small, around or less than $d = 0.2$ (Karchach & Verhaeghen, 2014). To be able to capture such small effect, the sample size needs to be increased considerably.

The study did not find a correlation between the N-Back components and transfer measures. This may be also be due to small sample size and homogeneity of the study participants, which may have led to restricted range of measurement in the transfer tasks. To assess the training gains, learning curves were fitted to the group data. However, to find meaningful individual differences in the learning gains and its relation with transfer measures, learning curves would need to be fitted at the individual level. Given the small number of data points per subject, we were not able to reliably perform such analyses in the present sample.

Use of strategies and its relation with task gains and transfer effects warrants further research since there is evidence that increased and/or more efficient strategy use may relate to better training outcomes (Dunning & Holmes, 2014). Some of the strategies used in this study were similar to the ones researched in other studies (e.g. De Simoni & Bastian, 2018; Laine, Fellman, Waris & Nyman, 2018; Morrison, Rosenbaum, Fair & Chein, 2016). Participants in this study were asked about their use of specific strategies retrospectively at the end of the training session. It is possible that this may have led to memory errors in reporting the specific strategies they employed for each working memory training task. Also, the order in which the training tasks were administered to the participants may have affected the use of strategies. Participants may have continued using the same strategy that they employed for the first training task for all subsequent

tasks. In addition, it has been found that the level of detail provided in the use of strategies by the participants is related to training gains (Laine et al., 2018). Level of detail was not investigated here.

4.8 Summary

The study aimed to better understand the basic underlying mechanism of working memory training and examine the hypothesis that training in a specific working memory component can provide scaffolding for higher-order processes. It specifically looked at the role of training the updating process, using a modified N-Back task. Some (e.g., Friedman et al., 2006) consider the updating component a key determinant of fluid intelligence and a critical component in eliciting transfer effects (e.g., Heinzl et al., 2016). It was found that updating is more malleable to training compared to focus switching, and thus training protocols employing working memory updating tasks may find larger benefits compared to employing other components of working memory. However, the study did not find support for the hypothesis that updating is the crucial factor in eliciting generalizable transfer effects.

REFERENCES

- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs?. *Psychological Bulletin*, 131(1), 30.
- Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2014). Improving fluid intelligence with training on working memory: a meta-analysis. *Psychonomic Bulletin & Review*, 22(2), 366-377.
- Beck, S. J., Hanson, C. A., Puffenberger, S. S., Benninger, K. L., & Benninger, W. B. (2010). A controlled trial of working memory training for children and adolescents with ADHD. *Journal of Clinical Child & Adolescent Psychology*, 39(6), 825-836.
- Bopp, K. L., & Verhaeghen, P. (2018). Aging and n-Back Performance: A Meta-Analysis. *The Journals of Gerontology: Series B*.
- Buehner, M., Krumm, S., & Pick, M. (2005). Reasoning= working memory≠attention. *Intelligence*, 33(3), 251-272.
- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97(3), 404.
- Cattell, R. B. (1973). *Measuring intelligence with the Culture Fair tests*. Champaign, IL: Institute for Personality and Ability Testing.
- Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin & Review*, 17(2), 193-199.
- Chooi, W. T., & Thompson, L. A. (2012). Working memory training does not improve intelligence in healthy young adults. *Intelligence*, 40(6), 531-542.
- Colom, R., Quiroga, M. Á., Shih, P. C., Martínez, K., Burgaleta, M., Martínez-Molina, A., ... & Ramírez, I. (2010). Improvement in working memory is not related to increased intelligence scores. *Intelligence*, 38(5), 497-505.

- Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence*, 32(3), 277-296.
- Conway, A. R., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, 30(2), 163-183.
- Cowan, N. (2005). *Working memory capacity*. Hove, East Sussex, UK: Psychology Press.
- Dahlin, E., Neely, A. S., Larsson, A., Backman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science*, 320(5882), 1510-1512.
- De Simoni, C., & von Bastian, C. C. (2018). Working memory updating and binding training: Bayesian evidence supporting the absence of transfer. *Journal of Experimental Psychology: General*.
- Dunning, D. L., & Holmes, J. (2014). Does working memory training promote the use of strategies on untrained working memory tasks? *Memory & Cognition*, 42(6), 854-862. doi:10.3758/s13421-014-0410-5
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General*, 128(3), 309.
- Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., & Nyberg, L. (2015). Neurocognitive architecture of working memory. *Neuron*, 88(1), 33-46.
- Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, 17(2), 172-179.
- Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: evidence for a developmental cascade. *Psychological Science*, 7, 237-241.

- Gaschler, R., Progscha, J., Smallbone, K., Ram, N., & Bilalić, M. (2014). Playing off the curve-testing quantitative predictions of skill acquisition theories in development of chess performance. *Frontiers in Psychology*, 5, 923.
- Garavan, H. (1998). Serial attention within working memory. *Memory & Cognition*, 26(2), 263- 276.
- Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science*, 24, 2409–2419.
- Heinzel, S., Lorenz, R. C., Pelz, P., Heinz, A., Walter, H., Kathmann, N., ... & Stelzel, C. (2016). Neural correlates of training and transfer effects in working memory in older adults. *Neuroimage*, 134, 236-249.
- Heinzel, S., Schulte, S., Onken, J., Duong, Q. L., Riemer, T. G., Heinz, A., ... & Rapp, M. A. (2014). Working memory training improvements and gains in non-trained cognitive tasks in young and older adults. *Aging, Neuropsychology, and Cognition*, 21(2), 146-173.
- Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2008). Enrichment effects on adult cognitive development can the functional capacity of older adults be preserved and enhanced? *Psychological Science in the Public Interest*, 9(1), 1-65.
- Hossiep, R., Turck, D., & Hasella, M. (1999). *BOMAT-advanced: Bochumer Matrizentest*. Hogrefe.
- Houben, K., Nederkoorn, C., Wiers, R. W., & Jansen, A. (2011). Resisting temptation: decreasing alcohol-related affect and drinking behavior by training response inhibition. *Drug and Alcohol Dependence*, 116(1), 132-136.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*, 105(19), 6829-6833.
- Jaeggi, S. M., Buschkuhl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory & Cognition*, 42(3), 464-480.

- Jaeggi, S. M., Studer-Luethi, B., Buschkuhl, M., Su, Y. F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning—implications for training and transfer. *Intelligence*, 38(6), 625-635.
- Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, 86(2), 199-225.
- Kane, M. J., Hambrick, D. Z., & Conway, A. R. (2005). Working memory capacity and fluid intelligence are strongly related constructs: comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, 131, 66–71.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133(2), 189.
- Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science*, 12(6), 978-990.
- Karbach, J., & Verhaeghen, P. (2014). Making Working Memory Work A Meta-Analysis of Executive-Control and Working Memory Training in Older Adults. *Psychological Science*, 25(11), 2027-2037.
- Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, 14(7), 317-324.
- Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., ... & Westerberg, H. (2005). Computerized training of working memory in children with ADHD-a randomized, controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 44(2), 177-186.
- Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, 24(6), 781-791.
- Laine, M., Fellman, D., Waris, O., & Nyman, T. J. (2018). The early effects of external and internal strategies on working memory updating training. *Scientific reports*, 8(1), 4045.

- Lange, E. B., Cerella, J., & Verhaeghen, P. (2011). Ease of access to list items in short-term memory depends on the order of the recognition probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 608–620.
- Li, S. C., Schmiedek, F., Huxhold, O., Röcke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: practice gain, transfer, and maintenance. *Psychology and Aging*, 23(4), 731.
- Lilienthal, L., Tamez, E., Shelton, J. T., Myerson, J., & Hale, S. (2013). Dual n-back training increases the capacity of the focus of attention. *Psychonomic Bulletin & Review*, 20(1), 135-141.
- Logan, G. D. (1994). On the ability to inhibit thought and action: A user's guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.
- Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, 49(2), 270.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100.
- Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, 81, 111–121.
- Morrison, A. M., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review*, 18, 46-60.
- Morrison, A. B., Rosenbaum, G. M., Fair, D., & Chein, J. M. (2016). Variation in strategy use across measures of verbal working memory. *Memory & cognition*, 44(6), 922-936.
- Oberauer, K. (2003). Selective attention to elements in working memory. *Experimental Psychology*, 50(4), 257.

- Oberauer, K., Schulze, R., Wilhelm, O., & Suß, H. M. (2005). Working memory and intelligence—their correlation and their relation: A comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, 131, 61–65.
- Price, J. M., Colflesh, G. J., Cerella, J., & Verhaeghen, P. (2014). Making working memory work: The effects of extended practice on focus capacity and the processes of updating, forward access, and random access. *Acta Psychologica*, 148, 19-24.
- Raven, J. C. (1990). *Advanced Progressive Matrices: Sets I, II*. Oxford: Oxford University Press.
- Raven, J.C. (1960). *Standard Progressive Matrices: Sets A, B, C, D, & E*. London: H.K. Lewis & Co.
- Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., ... & Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *Journal of Experimental Psychology: General*, 142(2), 359.
- Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General*, 123(4), 374.
- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124(2), 207.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103(3), 403.
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective?. *Psychological Bulletin*, 138(4), 628.
- Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *The American Journal of Psychology*, 89, 669-679.
- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, 153, 652-654.

- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.
- Thompson, T. W., Waskom, M. L., Garel, K. L. A., Cardenas-Iniguez, C., Reynolds, G. O., Winter, R., ... & Gabrieli, J. D. (2013). Failure of working memory training to enhance cognition or intelligence. *PloS One*, 8(5), e63614.
- Verhaeghen, P., Cerella, J., & Basak, C. (2004). A working memory workout: how to expand the focus of serial attention from one to four items in 10 hours or less. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(6), 1322.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: Psychological Corp.
- Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors*, 5, 7–17.
- Zinke, K., Zeintl, M., Rose, N. S., Putzmann, J., Pydde, A., & Kliegel, M. (2014). Working memory training and transfer in older adults: effects of age, baseline performance, and training gains. *Developmental Psychology*, 50(1), 304.